Genetic Control of Root Development in Arabidopsis

SHEN, WEIRAN

Genetic Control of Root Development in Arabidopsis

Weiran Shen

Department of Biosciences

Durham University

Submitted for the degree of PhD

December 2018

Abstract

The development of the root system represents an important feature of plant development, as the root is the site of water and nutrient uptake, it anchors the plant in the soil, and is a site of interaction with other organisms, which may be pathogenic (such as nematodes) or beneficial (such as mycorrhiza). The model organism *Arabidopsis thaliana* is an excellent model to study the genetic basis of root development, and to study the control of gene expression in response to hormones and as mediators of hormone action. A key question in developmental biology is how genes and hormone signalling systems interact to control cell identity, cell division, cell patterning and cell differentiation. To address this, the work described in this thesis focused of two genes, previously identified as playing important roles in root development - the *POLARIS* (*PLS*) gene, which encodes a 36 amino acids peptide; and *MERISTEM-DEFECTIVE* (*MDF*), which encodes an SR protein. In the course of this work, information on the pathways influenced by each gene was determined using high resolution RNA profiling followed by bioinformatics analysis, and information was used to identify pathways in which the two genes are involved. The results obtained show that PLS is required for both correct ethylene signalling and, independently, auxin biosynthesis in response to ethylene; and PLS exerts its effect via control of the tryptophan-independent pathway for auxin biosynthesis. It is also shown that MDF is a likely splicing factor, required for the regulation of auxin pathway genes and transcription factors expressed in the root meristem via the control of alternative splicing - it likely controls the balance in the meristem between stem cell identity and differentiation. The results provide new insights into the genetic and molecular mechanisms by which these genes regulate hormone signalling pathways to in turn control the development of the Arabidopsis root.

Statement of authorship

I certify that all of the work described in this thesis is my own original research unless otherwise acknowledged in the text or by references, and has not been previously submitted for a degree in this or any other university.

Statement of copyright

Acknowledgement

I would like to give my greatest appreciation to Prof Keith Lindsey. Thank you for taking me on board five years ago, giving me the chance to learn. Thank you for the patience, for always being there for me in the dark days and bring me back to positive, for going extra and far to help me when I was most in need! I can't ask for more for a supervisor, not to mention the music, gin and tonics and all the witty banter.

A massive thank you to Dr Jen Topping, not only for all the knowledge and techniques you generously passed on to me, but also for the motherly care, warmly guiding me through the problems in life when I am overwhelmed. Thank you so much for the love from you and Keith. Your way of thinking and doing things will guide me a long way through life. You and Keith made Durham more homely to me than it already is!

A special thank you to Dr Junli Liu. Thank you for inspiring me in the first place when I showed interest in data and programming, and the long meetings where you painstakingly taught me the programs and formulae.

Big thank you to Big J and Anna. Thank you for patiently showing me all the important equipment and teaching me techniques throughout my lab years. I am lucky to have had you two around in the early days of my PhD, making the transition to a postgraduate researcher much easier.

Big Sam, I made it, we made it! Thank you for the encouragement when I needed it the most since foundation year! Thank you for being a good classmate, practical mate, project mate, lab mate, and such a good friend!

Thank you to Beatriz and Mark for being such amazing teammates when we did RNA sequencing first time together. Learning with you made it so much easier with so much more fun!

Thank you, Gillian, Vicky and David, for all your support on my project. Without you my data would not be nearly as good, I am very grateful!

A big special thank you to all the colleagues and friends in our Lab 1004. Flora, small Sam, Amy, Helen, Ahmed, Kat, Swarna, Devina, and Chunli, thank you for the friendship and support, and all the drinks and laughter, making the lab a lovely family!

Thank you, Peng, Min, Lin, Jingze, Bowen, Ning, and Xiaohu for the lunchtime laughter and weekend games. Thank you for filling my life in Durham with Chineseness, making home sickness much easier to cope with!

Thank you the LSE MPA pack, Alar, Jeff, Jay, Chelsea, Justin, Habiba, Katya, Claire, and of course, Jimmy! You guys helped me to edge more towards becoming a global citizen. Thank you for the late night drinks in Hampstead, the grilled prawns in the Villa, and the dumplings in Ruislip. Those are good memories!

Good morning Jungwu, Gemma, Marina, Alex, and Sammy! You are among the funniest people I know. Thank you for being such good friends, making life in London wonderful! I look forward to the Fat Crab meet-ups. Let's keep the *tanxiaofengsheng* going for the years to come.

Thank you, Yao, Yu, and Zimo for the guidance, support, trust, fun and joy you gave me. Yao you are the freest soul I know, your philosophy of life makes me a freer person!

I am deeply grateful to my Mum and Dad. Thank you for sacrificing so much to give me such a life changing opportunity to study aboard. The journey has been truly amazing, and I have learnt so much! Also big thank you to Mum and Dad from Yangzhou. Thank you for the trust, encouragement, and positivity, especially in my dark days. I am so lucky to have you in my life.

Thank you Liz our beloved cat. Thank you for the love and companionship.

Last but not least, Xinyue, I love you! You are the toughest and loveliest girl I ever knew! I feel very lucky to have you as such a solid teammate for this lifetime. You have dug out the better part of me that I never knew I had. Thank you for all your love and support over the past years while I am working on my work, as well as my problems. I simply could not finish this thesis without you, and I would not be nearly as confident without your trust and encouragement. We have finally made it!

List of Abbreviations

The standard scientific conventions for protein and gene naming have been followed: wild type genes and proteins are in capitals and mutants are denoted by lower case, gene names are italicized whereas protein names are not.

Standard scientific abbreviations have been used for units of weight, length, amount, molarity, temperature and time.

Standard chemical element symbols, nucleic acid and amino acid codes are used.

1⁄2 MS10	Half strength Murashige & Skoog media
ABA	Abscisic acid
ACC	1-aminocyclopropane-1-carboxylic acid
Col-0	Columbia ecotype
CSLM	Confocal laser scanning microscopy
DAG	Days after germination
DEGs	Differentially expressed genes
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
GA	Gibberellic acid
GFP	Green fluorescent protein
GO	Gene ontology
GUS	β-glucuronidase
IAA	Indole-3-acetic acid
JA	Jasmonic acid
Log_2fc	Log_2 fold change NGS next generation sequencing
NPA	N-1-naphthylphthalamic acid
PAT	Polar auxin transport
PCR	Polymerase chain reaction
QC	Quiescent centre
qRT-PCR	Quantitative real time PCR
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
ROS	Reactive oxygen species

RSA	Root system architecture
WT	Wild type
YFP	Yellow fluorescent protein

Table of Contents

Chapter 1. Introduction 11
1.1 Study of plant development in *Arabidopsis thaliana* 11
1.2 Arabidopsis root and meristem development 12
The well characterised structure helps us to study how the processes of division, differentiation, and elongation are regulated, which involves complex interactions between plant hormones 13
1.3 Plant hormones interact with each other 13
1.4 Many auxin biosynthesis pathways were postulated 15
1.4.1 Trp independent pathway 15
1.4.2 Trp dependent pathways 16
1.5 Role of POLARIS 17
1.6 Role of MDF 18
1.7 Aims and objectives of the thesis 19
Chapter 2. Materials and Methods 20
2.1 Materials 20
2.2 Plant tissue culture 20
2.2.1 Seed sterilisation 20
2.2.2 Culturing Medium 21
2.3 Nucleotide extraction 21
2.4 Nucleotides Quantification and Quality Control 23
2.4.1 NanoDrop 23
2.4.2 TapeStation 23
2.5 Polymerase Chain Reaction (PCR) 23
2.5.1 Primers 23
2.5.1 cDNA synthesis 24
2.5.2 Standard PCR 24
2.5.3 Real time PCR (or Quantitative PCR, qPCR) 25
2.5.3 Gel electrophoresis 27
2.6 Gus staining and analysis 27
2.7 Imaging 28
2.8 RNA Sequencing 28
Principle 28
Chapter 3. Function and Signalling Relationships of POLARIS 34
3. 1 Introduction 34
3.2 Next-Gen RNA Sequencing 35

3.2.1 Quality Control and Data Integrity 35
3.3.2 Changes in gene expression in the *pls* mutant 48
3.4 qPCR validation of RNA sequencing data 52
3.5 Investigation of a role for PLS in auxin biosynthesis or transport 53
Summary 58
Chapter 4. Function and Relationships of MDF 59
4.1 Introduction 59
4.2 Next-Gen RNA sequencing on MDF 59
4.2.1 Quality control. 60
4.2.2 Analysis of differentially expressed genes using RNA-Seq Data 60
4.2.3 Overview of Differentially Expressed Genes using Gene Ontology (GO) Analysis . 62
4.2.*4* Key findings from the gene ontology analysis 67
4.2.5 Alternative Splicing analysis. 70
4.3 Bioinformatic study on role of MDF on response to stress 74
Summary 78
Chapter 5 The IAA Biosynthetic Network 79
5.1 Introduction 79
5.2 Auxin Biosynthetic Pathways 82
5.3 Construction of the IAA Biosynthesis Network 84
Summary 89
Chapter 6. Discussion 90
6.1 The POLARIS peptide is required for ethylene signalling control 90
6.2 The POLARIS peptide has a role in regulating photosynthesis genes 91
6.3 The POLARIS peptide is essential for ethylene-induced auxin biosynthesis in the Arabidopsis root tip 91
6.4 The IAOX pathway is likely to be the main POLARIS-regulated auxin biosynthesis pathway in Arabidopsis root tip 93
6.5 MDF is required for correct RNA splicing 95
6.6 MDF is required for correct auxin patterning and meristem development 95
6.7 MDF has a role in regulating responses to environment stresses. 96
6.8 Future work 97
6.8.1 ABS molecule in kinetic model 97
6.8.2 Confirming the IAOX pathway by measuring concentration of IAOX and glucosinolate in *pls* mutant 97
6.8.3 PLS and Photosynthesis 98
6.8.4 Validation of abnormal splicing events discovered in RNA sequencing experiment. 98

6.9 Summary 99
Bibliography 100
Appendices 107
Appendix 1 List of enriched gene ontology terms for up regulated genes in *pls* mutant. 107
Appendix 2 List of enriched gene ontology terms for down regulated genes in pls mutant. 110
Appendix 3 List of enriched gene ontology terms for up regulated genes in PLSox over expressor. 112
Appendix 4 List of enriched gene ontology terms for down regulated genes in PLSox over expressor. 114
Appendix 5 List of enriched gene ontology terms for up regulated genes in *mdf-1* mutant. 121
Appendix 6 List of enriched gene ontology terms for down regulated genes in *mdf-1* mutant. 126
Appendix 7 List of enriched gene ontology terms for up regulated genes in *mdf-2* mutant. 129
Appendix 8 List of enriched gene ontology terms for down regulated genes in *mdf-2* mutant. 135
Appendix 9 List of enriched gene ontology terms for differentially spliced genes in *mdf-1* mutant. 140
Appendix 10 Primer List 146

Chapter 1. Introduction

1.1 Study of plant development in *Arabidopsis thaliana*

The growth and development of plants require many plant hormones, or phytohormones, including auxin, ethylene, cytokinin, gibberellin, abscisic acid, jasmonic acid, salicylic acid and brassinosteroids, and they are essential in regulating many biological functions in response to stresses, pathogens, and other external stimuli.

Phytohormones have been studied for over a hundred years, and there has been significant progress made in recent years in our understanding of them in terms of biosynthesis and signalling pathways. It is also clear that many hormones work together, either synergistically or antagonistically, to regulate many aspects of plant development. Advances in technology and the use of computational techniques over past few years has allowed us to dissect the interactions systematically, leading to our lab's previous work on the kinetic model of a hormone crosstalk network (Liu et al., 2010). By incorporating key genes and hormones into the model, we examine how these key factors work with each other in the network in response to many external disturbances. The computational modelling approach has now been adopted widely, and provides a theoretical framework upon which to develop predictions that can then be tested experimentally (Moore et al., 2015).

Within this hormone crosstalk network, POLARIS (PLS), a peptide composed of just 36 amino acids, was found to be a key part in hormone signalling in Arabidopsis (Casson et al., 2002, Chilley et al., 2006, Liu et al., 2010). There is limited knowledge in the regulatory properties of small peptides in plants, but most are believed to act as ligands for receptors, such as in the CLAVATA network that regulates the size of the shoot apical meristem (Farrokhi et al., 2008). Previous work has confirmed that PLS may function as a metallochaperone for copper (I) ions, which are required for correct functioning of ethylene signalling (Mudge, 2016). The hormone crosstalk model predicts that there are roles for PLS in other parts of the network, especially in the biosynthesis of auxin. However, the specific mechanisms behind it are still to be discovered.

1.2 Arabidopsis root and meristem development

Arabidopsis thaliana is a small dicot flowering plant. It has been widely used as a model organism for plant sciences for over 100 years due to its small diploid genome, short life cycle, easy to care, and diverse natural variability. These traits make it a great model plant especially for genetics study (Somerville and Koornneef, 2002). In early 2000s, *Arabidopsis thaliana* became the world's first plant species to have its whole genome sequenced (Arabidopsis Genome, 2000). Its root is well characterised to have highly ordered structure, which is ideal for our study on hormone interactions and development (Figure 1) (De Smet et al., 2015).

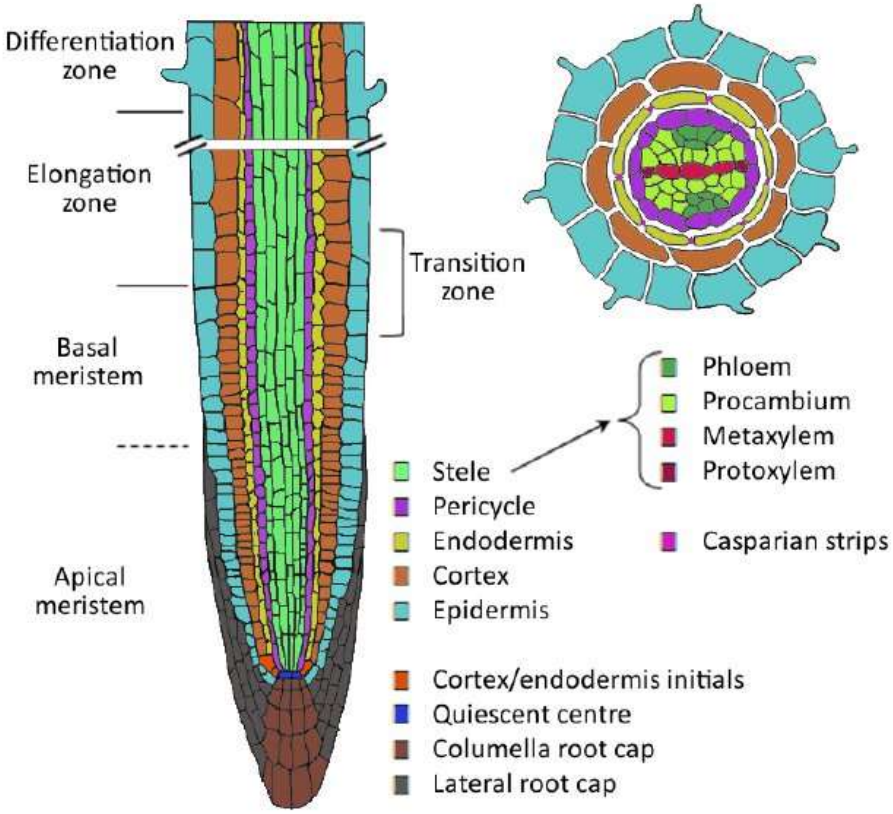


Figure 1. Organization of the Arabidopsis root (De Smet et al., 2015). (Left): Longitudinal section through the root showing apical-basal polarity. Different cell types (each differently coloured) are arranged in cell files, forming concentric single-celled layers surrounding the central vascular tissue. Distinct developmental zones are formed along the growing root. Cell division occurs in the meristematic zone, especially the apical meristem. Cell division rate slows down in the basal meristem and cells start to elongate in the elongation zone. The boundary between meristematic and elongation zone is indicated as the transition zone. Cell differentiation occurs in the differentiation zone; (Right): Radial polarity in on a cross section of the differentiated root zone showing the formation of root hairs and Casparian strips.

The primary root of *Arabidopsis thaliana* is made up of rings of distinctive cell files as shown in Figure 1. Each cell file maintains their specific structure and functions, and they all begin

from the stem cell niche, comprising the quiescent centre (QC) and the surrounding stem cells ('initials') in the root tip. The QC is composed of four cells which divide occasionally, regulating and maintaining the undifferentiated states of neighbouring stem cells (van den Berg et al. 2007). In each cell file in the meristem region, cells originated from the stem cells actively divide, pushing cells into the elongation and differentiation zones where cells elongate and differentiate into mature cells.

The well characterised structure helps us to study how the processes of division, differentiation, and elongation are regulated, which involves complex interactions between plant hormones

1.3 Plant hormones interact with each other

Hormones tightly regulate the growth and development of plants. Among the key plant hormones that attract most attention, auxin is the first to be discovered. Carrying the literal meaning of growth in Greek, auxin plays crucial roles in the development and growth plants. Within the complex network of hormone interactions, auxin controls fundamental cell activities, including elongation, division, and differentiation, all of which, in turn, will result in higher level development and growth, including lateral root formation, tropism and flowering etc (Davies, 2010). Despite all the attentions rightly deserved by auxin, we are still far from fully understanding how it is synthesized in plants, due to many postulated pathways, and possible redundent genes involved (Tivendale et al., 2014). In addition, there is also much more to learn about how auxin interacts with other genes and hormones to carry out its regulatory activity (Liu et al., 2010). This project tries to seek some light for both questions.

In the era of molecular biology, and facilitated by mutational studies using Arabidopsis, much progress has been made in finding the molecular mechanisms underpinning the pathways of hormone biosynthesis and signalling. Knowing that for these plant hormones, their activities are closely related depending on the developmental stage of the plant and its environmental conditions, it has proven to be a challenge to understand these interactions between these hormones and relevant genes. In 2010, Liu et al. (2010) published a model of the crosstalk between auxin, ethylene, cytokinin and *POLARIS* gene, which qualitatively matched relevant experimental data known so far (Liu et al., 2010). By connecting the interactions of phytohormones into a network, the systematic study has shed light on the

research on the entangled interaction network in a more structured way, while more experimental knowledge and parameterization are required to develop the model further.

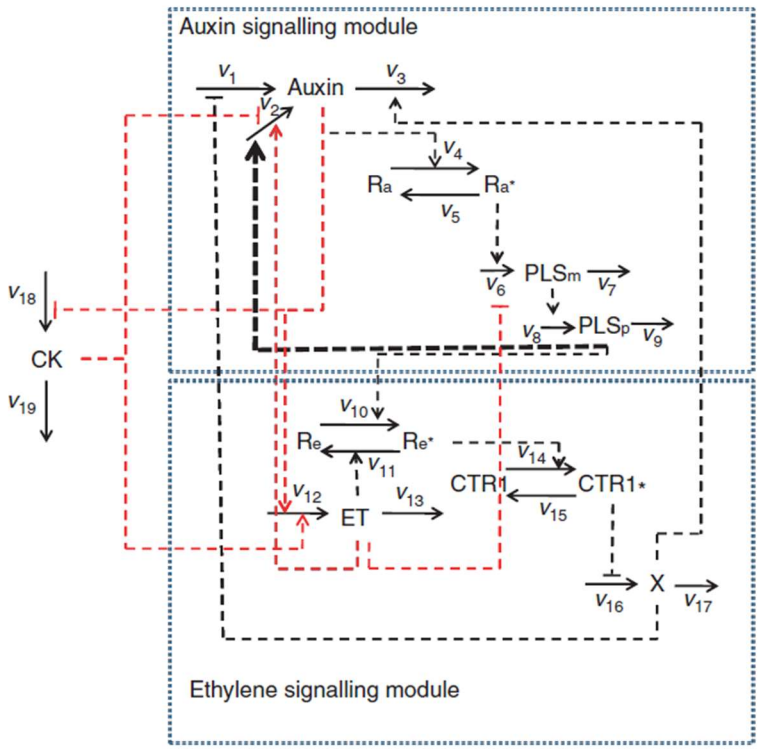


Figure 2. The original phytohormone crosstalk network (Liu et al., 2010). This shows where rate constants (*V*...) have been determined for interactions between auxin, ethylene (ET), the ethylene signalling kinase CTR1, cytokinins (CK) and POLARIS (PLS).

The model contains four major components: auxin, cytokinin, ethylene and PLS peptide, with particular emphasis on ethylene signalling and its interaction with PLS (Liu et al., 2010). Building the model was based on existing knowledge using mathematical modelling software, and kinetic parameters were tested and set to match available experimental data and the model was represented qualitatively. Investigating any component in the hormone network individually could be confusing without taking other affecting factors into account. Using molecular experiments with a complementing hormone crosstalk model provides a valuable approach to understanding the underlying interactions within the system as a whole. New components are being added to gradually develop the model. Not long after the model was published in 2010, the auxin efflux carrier PIN proteins were added to the model to include the transportation of auxin in the system, while also showing their interaction with PLS and other components in the model (Liu et al., 2013b). The PIN proteins comprise a family of 8 members, 5 of which are localised to the plasma membrane and mediate the

directional efflux of auxin, to establish concentration gradients; the remaining members are localized to the endoplasmic reticulum (ER), and their functions are not yet clear (Adamowski and Friml 2015).

1.4 Many auxin biosynthesis pathways were postulated

1.4.1 Trp independent pathway

Auxin is an important plant hormone involved in the regulation of growth and development of plants at every stage of the life-cycle (Woodward and Bartel, 2005). Although the study of its signalling pathway has proceeded quite well over last century, the synthesis of indole-3-acetic acid (IAA, auxin's predominant biologically active form) and its interaction with developmental and environmental factors are still poorly understood. Up until now, two main IAA biosynthesis pathways have been suggested, the tryptophan (Trp)-dependent and Trp-independent pathways (Woodward and Bartel, 2005, Normanly, 2010, Mano and Nemoto, 2012). They share the first few steps until indole synthesis, which is then converted to tryptophan in the Trp-dependent pathway (Figure 3). How the Trp-independent pathways utilize indole-3-glycerol phosphate or indole still remains unclear (Jian et al., 2000, Zhang et al., 2008).

Tryptophan is synthesized from chorismate in the chloroplast. A study on ethylene responses has uncovered an essential link point between ethylene and auxin activities (Stepanova et al., 2005). The inhibition of seeding root growth by ethylene is mediated by the products of WEI2 and WEI7 (WEAK ETHYLENE INSENSITIVE), which encode the two subunits of a tryptophan biosynthesis rate-limiting enzyme, anthranilate synthase that converts chorismate into anthranilate, an intermediate in the tryptophan synthesis pathway (Bohlmann et al., 1996). The *wei2 wei7* double mutant of Arabidopsis results in ethylene insensitivity in the root, which can be rescued by exogenous tryptophan or IAA (Stepanova et al., 2005). This agrees with previous evidence that auxin response is downstream of ethylene signalling and the *WEI2*, *WEI7* genes encode an essential part of the auxin biosynthesis pathway. Other relevant genes include *TSA1* and *TSB1* that encode α and β subunits of tryptophan synthase. *Arabidopsis* has another backup gene for the β subunit, *TSB2*, which is functionally redundant to *TSB1* (Last et al., 1991).

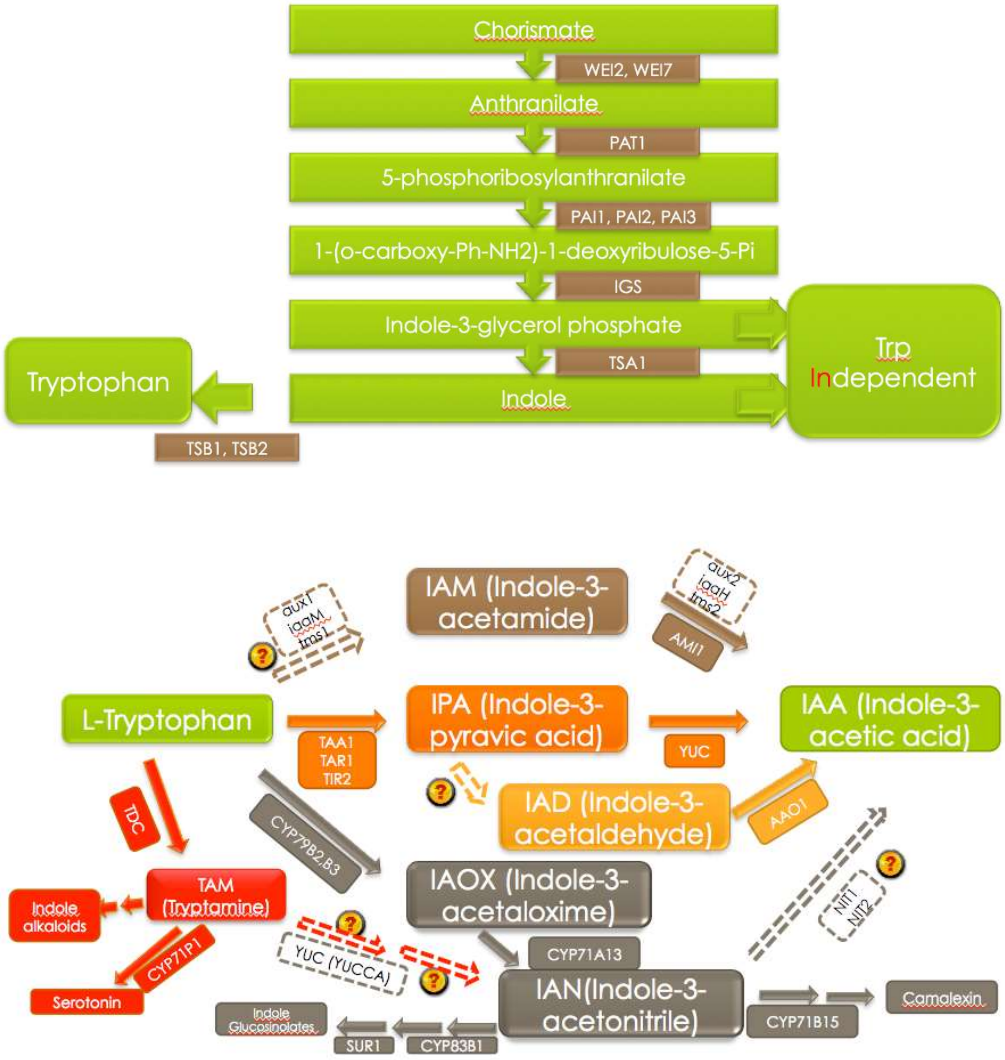


Figure 3. Postulated auxin biosynthetic pathways. (Upper Panel): The suggested tryptophan synthetic pathway. Brown boxes indicate the genes related to each step. **(Lower Panel):** The postulated Trp dependent pathways. Solid arrows and boxes indicate the genes, or their enzymatic activities that are known. Dashed arrows and boxes are genes or steps that remain unknown or poorly understood.

1.4.2 Trp dependent pathways

For Trp dependent auxin biosynthesis, many pathways have been suggested (Figure 3), including: the indole-3-acetamide pathway (IAM), the indole-3-pyruvic acid pathway (IPA), the tryptamine pathway (TAM), and the Brassicaceae species specific indole-3-acetaldoxime pathway (IAOX) (Chandler, 2009, Mano et al., 2010, Normanly, 2010, Mano and Nemoto, 2012, Woodward and Bartel, 2005). Despite being the first plant hormone discovered, there is still incomplete knowledge in the biochemistry of auxin biosynthesis. Among the many

postulated Trp dependent pathways, none has been proven indispensable, but it is suggested that the IAM and IPA pathways are the main routes to auxin synthesis.

1.5 Role of POLARIS

Previous work shows that POLARIS (PLS) plays an essential role in hormone signalling (Casson et al., 2002, Chilley et al., 2006, Liu et al., 2013a, Liu et al., 2010). PLS was identified in a promoter trap screen to identify genes expressed in the developing embryo of Arabidopsis, and the loss-of-function *pls* mutation exhibits defective root growth and hormonal responses (Casson et al. 2002). The *pls* mutant seedlings of Arabidopsis exhibit phenotypes of strong ethylene signalling without a noticeable increase in ethylene evolution, indicative of a role in ethylene signalling rather than synthesis (Chilley et al. 2006). Evidence shows that it regulates ethylene signalling by acting as a metallochaperone transporting copper (I) ion for ETR1, an ethylene receptor (Mudge, 2016). ETR1 is a member of a receptor family in Arabidopsis and other species, members of which form dimers containing a Cu(I) ion that is required for ethylene binding and signal transduction. The prediction is that the lack of PLS function would lead to enhanced ethylene responses as a consequence of receptor loss of function through lack of Cu(I) availability, as seen in *ran1* loss of function mutants - RAN1 delivers copper to the ER (Binder et al. 2010), and it is proposed PLS (which is also localised to the ER) retrieves this Cu and delivers it to the receptor complex.

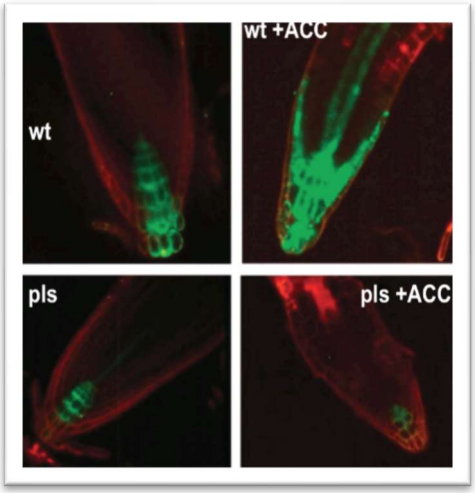


Figure 4. DR5:GFP showing auxin responses to exogeneous ethylene in wild type and *pls*. ACC treated seedlings are grown in the presence of 10 µM ACC added in the ½ MS10 growth medium. (Chilley et al., 2006)

Experimental data also suggest that there are direct interaction between POLARIS and auxin biosynthesis in response to ethylene. Ethylene plays a key role in many aspects of plant development, from fruit ripening, senescence and pathogenesis to the control of root growth and development (Li and Guo 2007). In part this is mediated through interactions with other hormones, and notably with auxin (Lewis et al. 2011; van der Poel et al. 2015). It is synthesised from methionine via the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC), which is often used experimentally to mimic ethylene gas treatment of plants (Bleecker and Kende, 2000).Figure 4 shows ethylene- (ACC-) induced auxin biosynthesis in wildtype, seen as DR5:GFP expression (Chilley et al., 2006). In the *pls* mutant, ACC does not induce auxin responses, showing that the PLS peptide is required for ethylene-mediated auxin responses. In other words, PLS peptide is acting upstream of the auxin response pathway, while down stream of ethylene signalling. One of the objectives of this project is to determine how PLS regulates auxin response pathways.

1.6 Role of MDF

Previous work has shown that the *MERISTEM-DEFECTIVE* (*MDF*) gene also plays a role in regulating auxin content and distribution in the Arabidopsis root (Casson et al. 2009). Its human homologue, the hSART-1 protein, is known to be a crucial component in pre-mRNA alternative splicing activity, which is a common mechanism among eukaryotes with the effect to have a much greater diversity in transcriptome compared with genome (Makarova et al. 2001). More work is required to further understand how MDF functions in the regulation of alternative splicing, and its effect on auxin in Arabidopsis.

Alternative splicing is a common phenomenon in eukaryotes, where pre-mRNA strands are spliced differently, promoting the diversity of transcriptome significantly. Previous work suggests that MDF might be a key component in alternative splicing apparatus, controlling expression level of key genes in correct meristem development and function, including auxin homeostasis (Casson et al., 2009). More work needs to be done to elucidate the mechanism in which it interacts with other hormones in the system.

1.7 Aims and objectives of the thesis

The overall aim of this project is to investigate the auxin biosynthesis and response pathways in *Arabidopsis thaliana*, to expand our knowledge in plant hormone crosstalk, especially how POLARIS and MDF proteins regulate and affect auxin distribution in the Arabidopsis root system.

The study starts from transcriptome profiling on multiple genotypes of Arabidopsis using the most recent technology in RNA sequencing. This gives an overview of how the mutations affect the expression level of all genes expressed in seedlings. Further investigation using bioinformatics techniques was planned to identify key biological functions or pathways affected by mutating *MDF* or *PLS* genes.

This led to experiments to investigate the relative contribution of auxin biosynthesis and transport in determine Arabidopsis root responses to ethylene.

To better understand the control of auxin biosynthesis pathways, a comprehensive network for auxin biosynthesis pathways was to be built using data from experiments and the wider literature.

Chapter 2. Materials and Methods

2.1 Materials

All chemical supplies are from Sigma (Poole, UK) or Fisher Scientific ltd (Loughborough, UK) unless otherwise stated.

Plant Material: All *Arabidopsis thaliana* wild type (Col-0 and C24) and mutant seeds are obtained from lab stocks unless otherwise stated.

Genotypes used in tissue culture:

C24 based lines: C24, *pls*.

Col-0 based lines: Col-0, PLSox, *etr-1*, *pls x etr-1*, *mdf-1, mdf-2,* MDFox.

All seeds used in this project are from lab stock. *pls* mutant seeds are from C24 background. All other genotypes are from Col-0 background. To minimise the adverse effect of noise raised from different background, all such experiments used wildtype from both backgrounds as control group.

The *pls* mutant, having a promoter trap T-DNA insertion into its short open reading frame, was first discovered in a promoter trap screening experiment (Casson et al., 2002).The PLSox overexpressing transgenic line was produced by incorporation 35S promoter of Cauliflower mosaic virus with PLS cDNA (Casson et al., 2002).

2.2 Plant tissue culture

2.2.1 Seed sterilisation

Sterilization is required prior to growing plant on nutrient rich media to prevent any fungi or bacteria growth.

It was done by submerging seeds in 70% ethanol v/v momentarily in 1.5ml Eppendorf tubes to remove the wax layer over the seeds, then the ethanol was discarded and replaced by 10% bleach v/v diluted from concentrated bleach (Tesco, UK) and left at room temperature for 10 minutes to ensure sufficient penetration. The seeds were then rinsed 5 times using sterile distilled water to remove residual bleach before stratified in the dark at 4°C for 4-7 days to synchronise germination.

2.2.2 Culturing Medium

The base culturing medium (1/2 MS10) is half strength Murashige and Skoog medium mixed with 10g/L sucrose and 2g/L Phytagel.

For every litre of culturing medium (1/2 MS10):

2.2g (half strength) Murashige and Skoog medium, 10g sucrose and 2g phytagel were mixed with distilled water with pH controlled to 5.7 with HCl and KOH, before autoclaving at 121 °C for 20 minutes.

2.3 Nucleotide extraction

SIGMA® Spectrum™ Plant Total RNA (Sigma-Aldrich, Gillingham, UK) kit was used for extracting total RNA from a small amount of plant tissue (~100mg), which were frozen in liquid nitrogen prior to extraction. The supplied protocol was strictly followed.

2.3.1 Grinding and lysing

About 100mg frozen plant tissue sample in an Eppendorf tube containing 500µl of lysis buffer (made by adding 10µl of 2-mercaptoethanol into lysis solution for every 1ml of lysis buffer) was ground on dry ice to fine powder with a pestle and incubated at 56°C for 2 minutes.

2.3.2 Filtration

The Sample was centrifuged at 14,000g for 3 minutes and supernatant was then transferred into a Filtration Column seated in a 2-ml collection tube before centrifuged again for 1 minute.

2.3.3 Binding RNA to Column

Filtration column was discarded and the clarified lysate was mixed with 500µl of binding solution before briefly vortexed. 700µl of the mixture was then transferred into a binding column seated in a new 2ml collection tube to be centrifuged at 14,000g for 1 minute. The flow-through liquid was discarded before the remaining mixture was transferred to the same biding column to repeat the same process. Most RNA should be bound to the column.

2.3.4 On-column DNase digestion

The column was washed by pipetting 300µl of Wash Solution I into the Binding Column and Centrifuge at 14,000g for 1 minutes. The Flow-through liquid was discarded from the tube and Binding Column was put back.

The DNase digestion mixture was prepared by mixing 10µl of DNase I (Catalogue No. D2816) with 70µl of DNase Digestion buffer (Catalogue No. D1566) gently by pipetting. Vortex was not used due to sensitivity of DNase I to physical denaturation.

With care, 80µl of the mixture was pipetted directly onto the centre of the filter in the Binding Column to incubate at room temperature for 15 minutes.

2.3.5 Three column washes and drying

After incubation, the column was washed again with 500µl of Wash Solution 1 before being washed twice with 500µl of Wash Solution 2 with the same process in the first wash in the beginning of DNase digestion step. With all the flow-through liquid discarded, the column was put back into the tube and centrifuged at 14,000g for 1 minute to dry.

2.3.6 Elution and collection

The column was then transferred to a new, clean 2ml Collection Tube before 50µl of solution was directly transferred onto the centre of the binding matrix inside the column to sit for 1 minute. Then the column was centrifuged at 14,000g for 1 minute. Purified RNA was then collected in the flow-through elute and ready for immediate use or storage at -80°C.

2.4 Nucleotides Quantification and Quality Control

2.4.1 NanoDrop

Measurements of nucleic acid concentration was carried out using a NanoDrop (Wilmington, Delaware, USA) ND-1000 Spectrophotometer, measuring absorption at 260nm, And NanoDrop ND-1000 V3.5.2 Software.

Absorption levels at 230nm and 280nm were also used as an indication to the purity of nucleic acid samples.

2.4.2 TapeStation

Agilent 2200 TapeStation was used to measure concentration of nucleic acid samples as well as their strand length distribution. **D1000 ScreenTape** (C/N 5067-5582) and **D1000 Reagents** (P/N 5067-5583) were used to run electrophoresis on DNA samples, while **High Sensitivity RNA ScreenTape** (C/N 5067-5579), **High Sensitivity RNA ScreenTape Sample Buffer** (C/N 5067-5580), and **High Sensitivity RNA ScreenTape Ladder** (C/N 5067-5581) were used for RNA samples.

2.5 Polymerase Chain Reaction (PCR)

2.5.1 Primers

The primes used in PCR reactions were designed using Primer-Blast (Http://www.ncbi.nlm.nih.gov/tools/primer-blast/) online tool and synthesised by MWG Eurofins (http://www.eurofinsdna.con/). Appendix 10 contains the full list of primers used in this project.

2.5.1 cDNA synthesis

cDNA was synthesized by reverse transcription reaction, where the Poly-A RNA was copied into its complementary DNA form. So, the cDNA can be used as template for amplification in PCR reactions. Super Script III reverse transcriptase form SIGMA was used for cDNA synthesis.

The following reaction mixture was added into a 20µl tube for first stage of incubation at 65°C for 5 minutes:

5ng Total RNA in water	10µl
10µM Oligo dT	1µl

Then the followings were added into the reaction mixture for the second stage of incubation at 50°C for 50 minutes:

10x FX buffer	2µl
50mM $MgCl_2$	2µl
0.1 DTT	2µl
RNaseOut (Recombinant Ribonuclease Inhibitor)	1µl

After the 2nd incubation, the temperature was increased to 85°C for a further 5 minutes to denature the enzymes in the mixture. The mixture was then put on ice for 1 minutes to cool down, before adding 1µl of RNase H for the last incubation at 37°C for 20 minutes to remove the RNA.

The final product was diluted 1 in 4 and stored at -20°C until needed for PCR.

2.5.2 Standard PCR

For standard PCR reactions, Taq DNA polymerase from Bioline was used along with supplied 10x reaction buffer and 50mM $MgCl_2$ solution.

Primers were ordered from MWG EUROFINS as freeze dried powder, which were re-suspended to desired concentration with Milli-Q ultrapure deionised sterile water.

The templates used were mainly cDNA synthesized from Poly(A) RNA.

A standard PCR reaction mix contains:

DNA template	10-200ng
Forward primer	0.2µM
Reverse primer	0.2µM
50mM $MgCl_2$	1.5µl
10x reaction buffer	5µl
dNTP mix	1mM
Taq DNA polymerase	2.5 units

The volume was made up to 50µl with Milli-Q water in a 0.5ml PCR tube before being placed in a Thermal Cycler.

A typical amplification program:

Denaturation at 94°C for 2 minutes, followed by:

30 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 72°C for 1 minutes,

Then a final extension of 10 minutes at 72°C.

The size and concentration PCR products were checked by running 10-20µl of reaction mix on an agarose gel (0.7/2% thickness depending on expected length of products).

2.5.3 Real time PCR (or Quantitative PCR, qPCR)

Compared to standard PCR, the real-time PCR can detect the amplified DNA as the reaction cycles in real time. It is achieved by using a fluorescent dye, SYBR, which dramatically increases its fluorescence when bound to double strand DNA. The fluorescence is monitored after each reaction cycle and fluorescence data is analysed. By comparing with positive and

negative control, a comparative expression level of a target gene can be accurately quantified.

Real time PCR experiments were performed using either SIGMA-ALDRICH SYBR® Green JumpStart™ Taq ReadyMix™ or BIOLINE SensiFAST™ SYBR® No-ROX Kit as SYBR mix.

For each reaction, following components were mixed into a 20µl cocktail:

SYBR mix	10µl
10µM forward primer	0.8µl
10µM reverse primer	0.8µl
Template	up to 8.4µl
Milli-Q water	as required

The cycling program was set as following:

Polymerase activation	95°C for 2 minutes
40 cycles of	
Denaturation	95°C for 5 seconds
Annealing	60-65°C for 10 seconds
(Annealing temperature is dependent on primers used)	
Extension	72°C for 20 seconds

Each real-time PCR experiment is carried out using tissue collected from at least three independent biological replicates, each of which also have three technical replicates. The data obtained from real-time PCR were analysed using the Rotorgene Q series software v1.7. Relative expression level of each gene was normalised using housekeeping gene for each sample, and analysed by comparative quantification method using an assumption free, linear-regression approach (Ramakers et al., 2003).

2.5.3 Gel electrophoresis

As DNA molecules are negatively charged and they travel at different rate in agarose gel depending on their size, electrophoresis can be used to separate DNA molecules in the PCR product by size.

The gel is made by mixing Agarose Multi-Purpose (Bioline) with 1x TAE buffer (diluted 1 in 10 from 10X TAE buffer: 242g Tris, 37.2g $Na_2EDTA{\cdot}2H_2O$, 57.1ml glacial acetic acid, and water in 5L total volume) and heated to 100°C using microwave to produce a 1% w/v mix. Ethidium bromide was mixed to the heated solution before it is poured into an appropriate mould where the gel sets. 5x DNA loading buffer (Bioline) was mixed with PCR product in 1:4 v/v ratio, before loaded in the wells in the gel, with a separate lane for a suitable Hyperladder as the standard to determine the size of each separated band. The gel was run at 80V for about 40 minutes before imaged using BioRad Gel-Doc 1000 (BioRad).

2.6 Gus staining and analysis

The GUS (beta-glucuronidase) reporter system is used to analyse promoter localisation and activity in transgenic plants. The expression of the GUS enzyme can be identified by incubation tissues in X-Gluc solution, which is converted to a blue precipitate by GUS enzyme in localised tissue.

The staining solution comprises 1mM N-N-dimethlfomamide in 100mM sodium phosphate (pH7.0), 10mM EDTA, 0.5mM potassium ferricyanide, and 0.1% v/v Triton X buffer (Topping and Lindsey, 1997). The chloral hydrate solution is made up with 8g chloral hydrate, 1ml glycerol and 2ml water.

To perform the staining process, individual seedlings were immersed in the staining solution in a 1.5mm Eppendorf tube. Optimum time for staining is determined by a time-course analysing the effect of the staining prior to the experiment. The staining process is stopped by replacing the staining solution with 98% ethanol, which also preserves the sample.

Before imaging, each seedling is rehydrated by replacing ethanol with water, then transferred to a slide with chloral hydrate solution.

2.7 Imaging

2.7.1 Compound Light microscope

A Zeiss Axioskop compound microscope (Carl Zeiss, Cambridge, UK), were used to examine histological tissue sections. Digital pictures were obtained with a QImiging Retiga-2000r camera (Photometrics, Marlow, UK) mounted on the microscope.

2.7.2 Confocal microscope

Tissues with florescent markers were examined using a Zeiss LSM 800 laser scanning microscope using either x40 or x63 oil immersion objectives. Roots were mounted on slides in sdH_2O, covered by a 1.5 x 1.5 mm cover slip. The parameters for the excitation of fluorophores were as following: GFP by 488 nm using Argon laser, and YFP by 514nm using Argon laser.

2.8 RNA Sequencing

-By Next Gen Illumina dye sequencing

Principle

The Illumina dye sequencing is a new generation of sequencing technique to determine the series of base pairs in DNA. It is achieved by synthesizing an identical copy of template ssDNA, which is fixed on a flow cell, with fluorescently labelled nucleotides with terminal 3' blocked, one base at a time, and a picture is taken after each base is added to the strand. Then the fluorescent dye, along with the terminal 3' blocker is chemically removed, to allow next cycle to begin. As the four nucleotides have different colours, the dots on each picture gives information on which base is added to a DNA strand at given cycle, hence revealing the sequence of a DNA segment.

With DNA sequencing technology, RNA sequencing was achieved by sequencing the cDNA synthesized from extracted mRNA.

2.8.1 Library preparation:

a. RNA extraction

As RNA sequencing requires high quality complete RNA samples, TRIzol Reagent was used in the lysis part of RNA extraction to extract as much RNA molecules as possible, especially the smaller ones, from plant tissue culture.

1. **Lyse samples and separate phases**

   200mg of each sample was wrapped in foil and frozen in liquid nitrogen, before being ground into fine powder in a frozen pestle and mortar. Ground tissue was collected into a 1.5ml Eppendorf tube before 1000µl TRIzol Reagent is added into the tube and mixed by inverting the tube a few times. After 5 minutes of incubation at room temperature to allow complete dissociation of the RNA from cells, 200µl chloroform (1:5 v/v to TRIzol used) was added to the tube, mixed by inverting a few times, before incubation at room temperature for 3 minutes before centrifuged at 12000g at 4°C for 15 minutes.

   After centrifuging, the upper clear aqueous phase containing the RNA was transferred into a new tube very carefully to avoid contamination from the interphase or organic layer in the tube.

2. **Isolate RNA**

   Protocol for SIGMA® Spectrum™ Plant Total RNA kit was used to isolate RNA from the “Binding RNA to Column” step.

3. **Quality Control**

Protocol for TapeStation was used to analyse the concentration and quality of isolated RNA samples.

b. Library Preparation

Illumina TruSeq Stranded Total RNA Sample Preparation Guide (Part # 15031048 Rev.E) was used for the Library preparation. It comprises the following processes:

Ribo-Zero Depletion
RNA fragmentation
First Strand cDNA synthesis
Second Strand cDNA synthesis
Adenylate 3' ends
Adaptor ligation
DNA Fragment Enrichment
Library validation
Normalisation and pooling

c. Sequencing

The pooled library sample was handed to sequencing lab for 125 base paired end sequencing on an Illumina HiSeq 2500 sequencer.

2.8.2 Data Processing

The raw data output from the sequencer is sequencing information for each of the billions of short DNA strands (up to 250 bps including adaptors and tags) present in the library pool. Therefore, adequate processing was required before the data can be used for analysis. Firstly, the sequencing

information from each sample needed to be separated from the pooled library. This was achieved by identifying the unique tags at either ends of each read, which were assigned during the library preparation process. Secondly, the reads needed to be compared against the sequence of a known genome as reference to be able to identify the location of each read in the reference genome. Thirdly, with its location data, each read was be identified as part of a gene by using a gene transfer format (GTF) file containing the coordinates of genes within the genome. Depending on research requirements, different quantification methods can be used to convert the reads into gene counts.

The following is a list of programmes used for data processing:

a. Basic genome alignment using the following programs to align the raw reads against TAIR10(EnsemblePlants) genome.
    i. TopHat (Controls alignment process)
    ii. Bowtie (aligns reads against genome)
    iii. Cufflinks (link mapped reads into transcripts)
b. Sequence file indexing
    i. Samtools (Indexes and sorts the binary sequence alignment files (BAM files) and convert them into readable (SAM) files)
c. Quantification
    i. CuffDiff2 (Differential expression analysis tool)
    ii. HTSeq (A Python package that counts reads)
    iii. DeSeq (As part of R based Bioconductor, it is a differential gene expression analysis based on the negative binomial distribution)

2.8.3 Data analysis

a. Differential gene expression analysis

Gene count, or transcription expression level, was estimated using GTF files that contain gene coordinate information within the genome. Various factors can affect the result of quantification, such as transcript length, total number of reads, and sequencing biases. It needs to be considered that longer genes would have more reads as a result of accumulating more fragments than smaller genes. This is not important when comparing the expression level of a certain gene across several different samples, but it is necessary for ranking the gene expression level within the same sample. Another challenge facing differential gene expression level analysis is the problem of many related transcripts' sharing a lot of their reads. Many programs developed sophisticated algorithms to tackle this problem. As an example, TopHat analyses the non-uniform read distribution along the gene length to allocate reads to different transcripts accordingly.

b. Alternative splicing analysis

We adopted the approach which detects the presence of expression isoforms by looking at reads spanning across exon junctions. Tools like DEXseq compare significant differences in read counts on exons between transcripts. Using the traditional classification method, there are 5 basic types of alternative splicing events:

1. Exon Skipping, where an entire exon is skipped or retained in a transcript.
2. Mutually exclusive exons, where one of a group of exons is included in transcript after splicing, but not more than one at the same time.
3. Alternative donor site, where part of an exon is skipped or retained before jumping to the next exon.
4. Alternative acceptor site, where part of an exon is skipped or retained when starting transcription at a new exon.

5. Intron retention, where the intron between two exons is retained or skipped.

DEXseq (http://bioconductor.org/packages/release/bioc/html/DEXSeq.html) is used for identifying differential exon usage between RNAseq samples.
DSGseq (http://bioinfo.au.tsinghua.edu.cn/software/DSGseq/) is used for comparing differentially spliced genes from two RNAseq samples.

Chapter 3. Function and Signalling Relationships of POLARIS

3. 1 Introduction

The aim of the experimental work in this chapter is to investigate the function of *POLARIS* and its relationships with other genes and hormones, and in particular whether it has a direct influence on auxin biosynthesis independent of ethylene signalling.

The *POLARIS* (*PLS*) gene was first identified in a mutagenesis screen, and was found to have a T-DNA containing a promoterless *gusA* gene inserted into a small open reading frame (ORF) of the gene, interrupting its expression (Topping et al., 1994). Having 108bp of coding sequence, *PLS* encodes a predicted 36-amino acid POLARIS (PLS) peptide (Topping et al., 1994). The phenotype of homozygous mutant seedlings includes a short primary root, reduced polar auxin transport, and low auxin accumulation. It was previously proposed that the peptide regulates auxin distribution in the root via effects on ethylene signalling (Chilley et al., 2006).

As the very first phytohormone to be discovered, auxin is involved in many aspects of plant development and growth, and it is part of a complex network where phytohormones and relevant genes interact with each other. Previous studies demonstrated that ethylene can induce auxin biosynthesis and transport in the root (Stepanova et al., 2007, Swarup et al., 2007, Stepanova et al., 2008). In wild type seedlings, applying the ethylene precursor ACC shows a dramatic effect in elevating auxin accumulation in the root meristem region. However, evidence shows that exogenous ACC treatment failed to rescue the suppression of auxin accumulation in the root tip in *pls* transgenic line (Liu et al., 2010), suggesting that PLS might have a role in auxin biosynthesis independent of ethylene signalling.

Using next generation RNA sequencing technology, I was able to compare the transcriptome of the *pls* mutant and transgenic overexpressing line with wild type. The objective was to use differential expression analysis to determine whether the ethylene phenotype of *pls* is also reflected in expression levels of ethylene-responsive genes, and whether this is linked to any upregulation of auxin biosynthetic genes. We can hypothesise that, if auxin biosynthetic genes are not up-regulated by the lack of PLS peptides in *pls* mutant, then PLS is required for the transcription of auxin biosynthesis-related genes. If such genes are, however, upregulated, this would suggest a post-transcriptional mechanism of PLS action, to account for the low auxin phenotype of the *pls* mutant.

3.2 Next-Gen RNA Sequencing

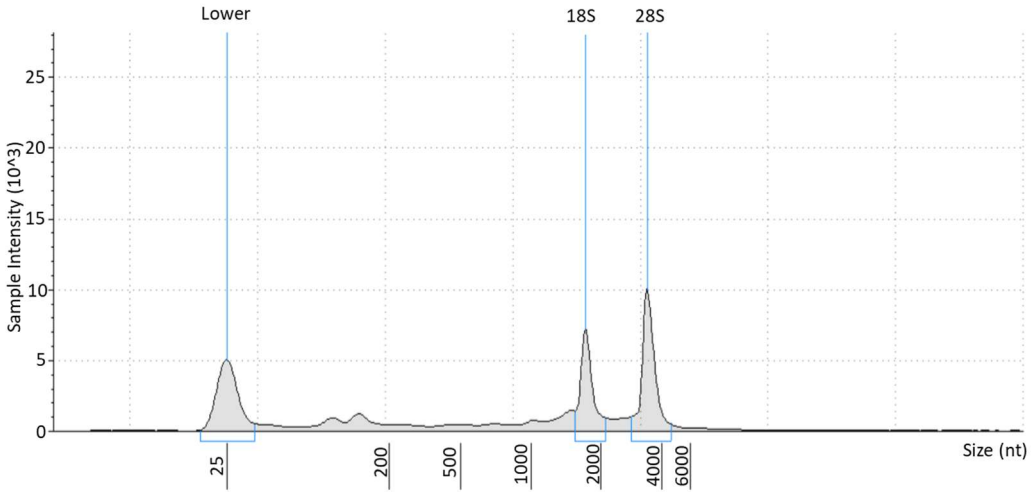
Thanks to the recent development in sequencing technology, the whole transcriptome of tissue samples can be used to give gene differential expression information between different genotypes or treatments. Using the Next Gen RNA sequencing platform from Illumina, we sequenced the RNA profile of C24, *pls*, Col-0 and *PLSox* genotypes.

3.2.1 Quality Control and Data Integrity

Three biological replicates were used in the RNA sequencing experiment for each of the C24, Col-0, pls, and PLSox genotypes. Their RNA quality and concentration were monitored throughout the experiment. Three critical points for quality control were 1) end of RNA extraction, 2) end of library preparation, and 3) end of sequencing run.

3.2.1.1 RNA Quantification and Quality Control

Total RNA was extracted from each sample, and was run on TapeStation to test their quality and integrity. All the samples in this experiment achieved satisfactory concentration and excellent RIN^{e} quality scores. Figure 3.1 shows a representative TapeStation result, and high quality RNA was confirmed by the clear presence of ribosomal RNA. The concentration of each sample was also confirmed by NanoDrop spectrometry analysis.



Peak Table

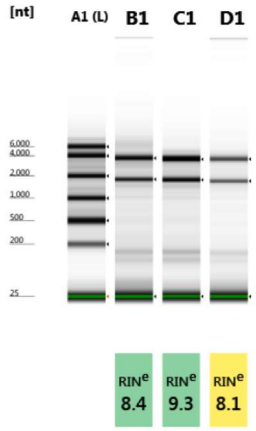
Size (nt)	Calibrated Conc. (ng/µl)	Assigned Conc. (ng/µl)	Peak Molarity (nmol/l)	% integrated Area	Peak Comment	Observations
25	47.5	47.5	5590	-		Lower Marker
1731	20.9	-	35.6	40.06		18S
3384	31.3	-	27.2	59.94		28S

RNA Data Table

28S/18S (Area)	1.4
Total RNA Area	3.59
rRNA Area	1.10

Figure 3.1 – Representative TapeStation electrophoretogram of an RNA sample used for RNA sequencing. Each sample was stained with fluorescent dye and separated on an electrophoresis tape. The x-axis shows the size of detected molecules, calculated from the length it travelled from the starting point, calibrated with a standard ladder. The y-axis shows fluorescence intensity. The bands in the gel image shows the intensity of each RNA constituents. Smaller molecules travel faster so they appear at the bottom of the gel and shown at left hand side of the graph. The two peaks on the right represent 18s and 25s subunits of ribosomal RNA, and the one on the left is the marker in the gel.

Gel Image



RNA
Contrast: 0.50

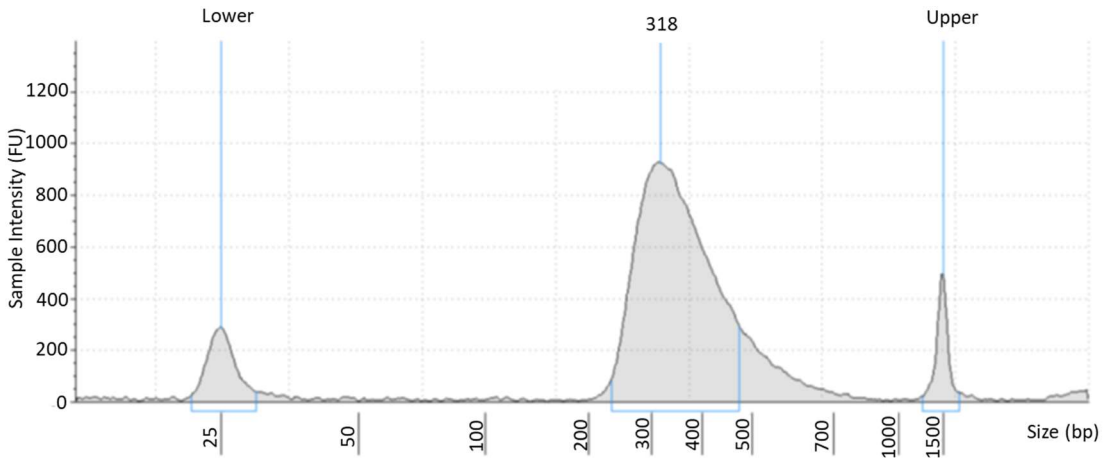
Sample Info

Well	RIN^e	28S/18S (Height)	28S/18S (Area)	Conc. [ng/µl]	Sample Description	Alert	Observations	Total RNA Area	rRNA Area
A1	-	-	-	133	Ladder		Ladder	-	-
B1	8.4	1.4	1.5	56.7	EM101 C			2.09	0.57
C1	9.3	1.4	1.6	66.2	C24 C			2.44	0.86
D1	8.1	1.2	1.3	30.5	PLS-ox C			1.12	0.33

Figure 3.2 – Representative TapeStation gel images of some RNA samples used for RNA sequencing. Each sample was stained with fluorescent dye and separated on an electrophoresis tape. From left to right, the first column is a standard ladder, showing nucleotide length in the scale on the left. The 3 columns next to the ladder are samples from *pls* (B1), C24 (C1), and *PLSox* (D1) respectively. The RIN^e score ranges from 0 to 10, measuring the integrity of RNA samples.

3.2.1.2 Library Quantification and Quality Control

To construct a transcriptome library for each sample, the total RNA went through a lengthy process, including removal of ribosomal RNA, cutting large molecules into the size, converting them to cDNA while adding adaptors and indexes, and controlled amplification using PCR. Figure 3.3 shows a representative Tape Station electrophoretogram of a Library sample used for RNA sequencing. All samples accumulated large amount of DNA molecules of ca. 300 bp, which is ideal for this 2-sided sequencing experiment.



Col-0 A

Peak Table

Size [bp]	Calibrated Conc. [pg/µl]	Assigned Conc. [pg/µl]	Peak Molarity [pmol/l]	% Integrated Area	Peak Comment	Observations
25	363	-	22400	-		Lower Marker
318	2990	-	14500	100.00		
1,500	250	250	256	-		Upper Marker

Gel Image

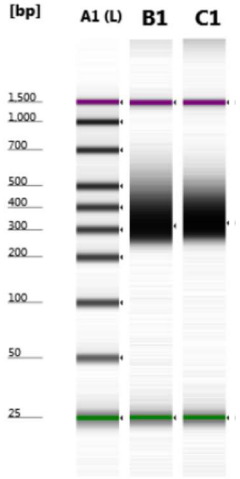


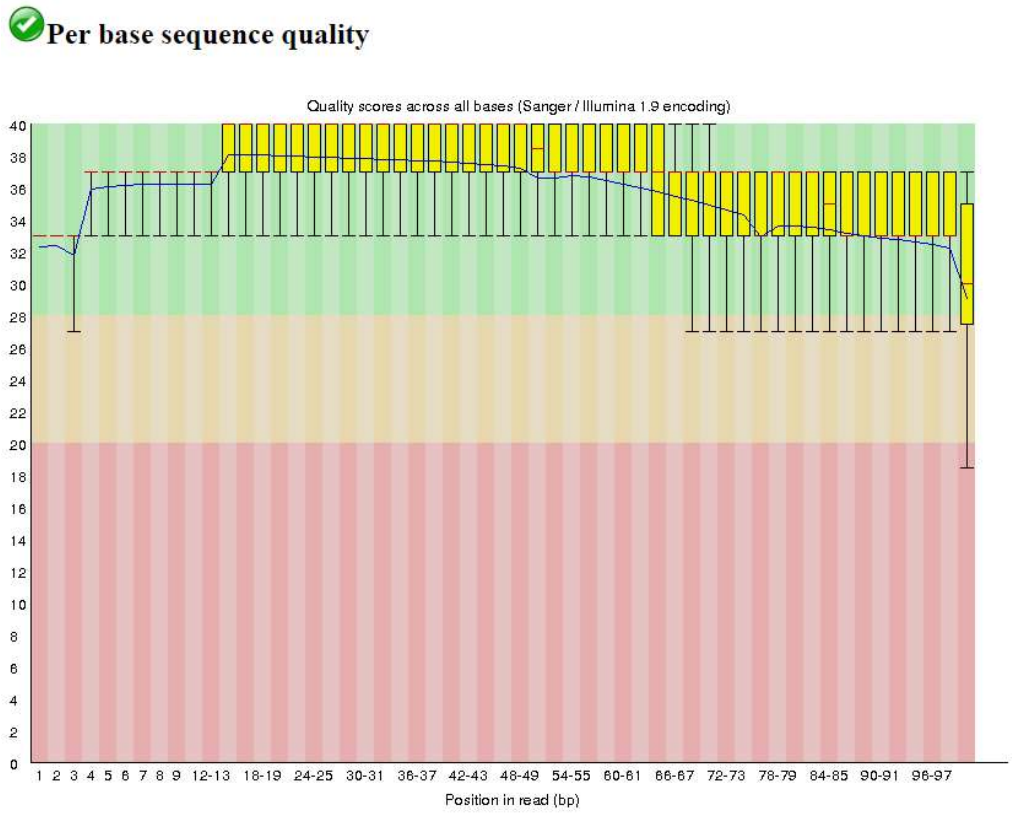
Figure 3.3 – Representative TapeStation electrophoretogram of an RNA sample used for RNA sequencing. Each sample was stained with fluorescent dye and separated on an electrophoresis tape. The x-axis shows the size of detected molecules, calculated from the length it travelled from the starting point, calibrated with a standard ladder. The y-axis shows fluorescence intensity. The bands in the gel image shows the intensity of each of the DNA constituents. Smaller molecules travel faster so they appear at the bottom of the gel and are shown at the left hand side of the graph. The large peak in the middle represents the DNA molecules with the desired size of around 300 bp. The two smaller peaks at either side of the graph are markers in the gel.

As the sequencer is very sensitive to library concentration, any variation between samples would be amplified in the sequencing process, making differential gene analysis more

difficult. It is important to normalise the concentration of libraries before pooling them together. To achieve accurate measurements, qPCR was used to determine the concentration of each library. The libraries of all samples were then pooled together and sent for Illumina sequencing.

3.2.1.3 Data Quality

Raw data from the sequencer was tested using FastQC software, which gives a score for each base in all reads. As shown in Figure 3.4, high quality data were obtained for all 12 samples according to FastQC read quality reports. The quality score of the majority of bases are well within the desired green zone (score 28-40) for this experiment. (Figure 3.4)



Genotype	Average Quality Score per Read
C24	37
EM101	37
Col-0	37
PLSox	37

Figure 3.4: The upper panel is a representative Per Base Sequence Quality report. The plot shows an overview of the quality scores across all bases at each position of all reads. The x-axis shows the position (bp) of a base in the read, and the y-axis shows the quality score. The red line is median value; the yellow box represents the inter-quartile range (25-75%); The upper and lower whisker indicate 10% and 90% points; The blue line is median score. The table below the graph shows the average quality score per read of each genotype tested, where all four genotypes achieved 37 out of 40, well over the threshold for best quality at 28.

The Illumina sequencer used in this experiment is capable of reading 125 bases in each direction in total. Accounting for the depleted 12 bp of adaptors and 12 bp of indexes, most reads in the output file have 101 bases. On most Next Gen sequencing platforms, the read quality degrades as the run progress (Manley et al., 2016). Therefore, the Per Base Read Quality plot (Figure 3.4) exhibits slightly decreasing average quality score along the reads. Quality warning is only issued when the median score (red bar) of any base is below 25. In more severe cases, the program will register failure when the median score hits below 20, or the lower quartile score (yellow box bottom) is below 5 (Andrew, 2016).

Due to the fact that the samples from all 4 genotypes were run together in the same lane, all successful and failed reads are uniformly distributed between them. As a result, their quality scores are highly similar to each other (Figure 3.4, lower panel).

3.3. Overall differential gene expression comparison between *pls* and PLSox

After passing the quality control, the raw data obtained from the sequencer subsequently went through the process of trimming off the adaptors, alignment to the known *Arabidopsis thaliana* genome (Tair10), and normalising and counting the genes. The result is a file containing the expression level of all genes in each sample.

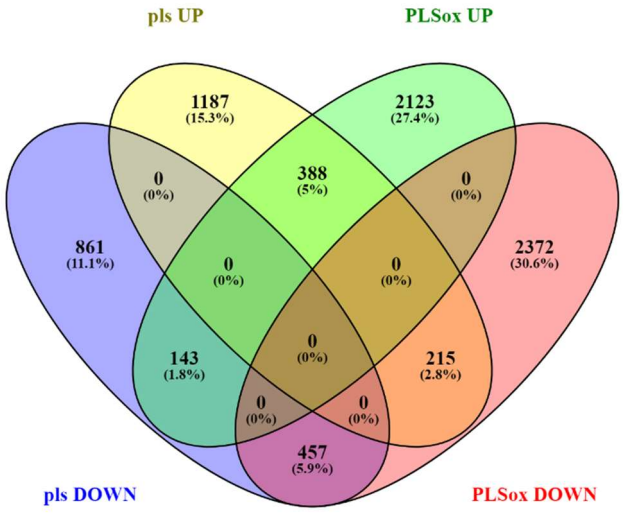


Figure 3.5 Venn diagram showing number of genes that have changed expression levels in the *pls* mutant and PLSox compared to wildtype ($p<0.1$). In total, expression of 3251 genes was found to be significantly different in *pls* whereas for PLSox the number is 5698.

3.3.1 Overview of Differentially Expressed Genes using Gene Ontology (GO) Analysis

High-throughput experiments can produce large amounts of data which are difficult to interpret. As a good example, the RNA sequencing experiment in this project identified thousands of differentially expressed genes (DEGs) from each sample. In order to interpret the data, gene ontology (GO) enrichment analysis was carried out using the online platform agriGO (http://systemsbiology.cau.edu.cn/agriGOv2/). This analysis identifies frequently occurring GO terms within the list of DEGs to show the proposed biological processes that are affected in the *pls* mutant and PLS overexpressor. However, the list of enriched GO terms can still be long and redundant, making it hard to interpret. To make the result more intelligible, the long list of enriched GO terms is further reduced using REVIGO (revigo.irb.hr), which groups GO terms semantically based on their functional similarity into clusters, using a single, most relevant GO term picked from each cluster as the representative for that cluster. To visualise the output, the clusters are displayed in a tree map as rectangles, which are grouped further into superclusters each assigned a colour, with the sizes of the rectangles represent the p-values of the enriched GO terms. This analysis gives an overview of the DEG profile of each sample, making it easier to identify the potential biological processes and pathways of interest.

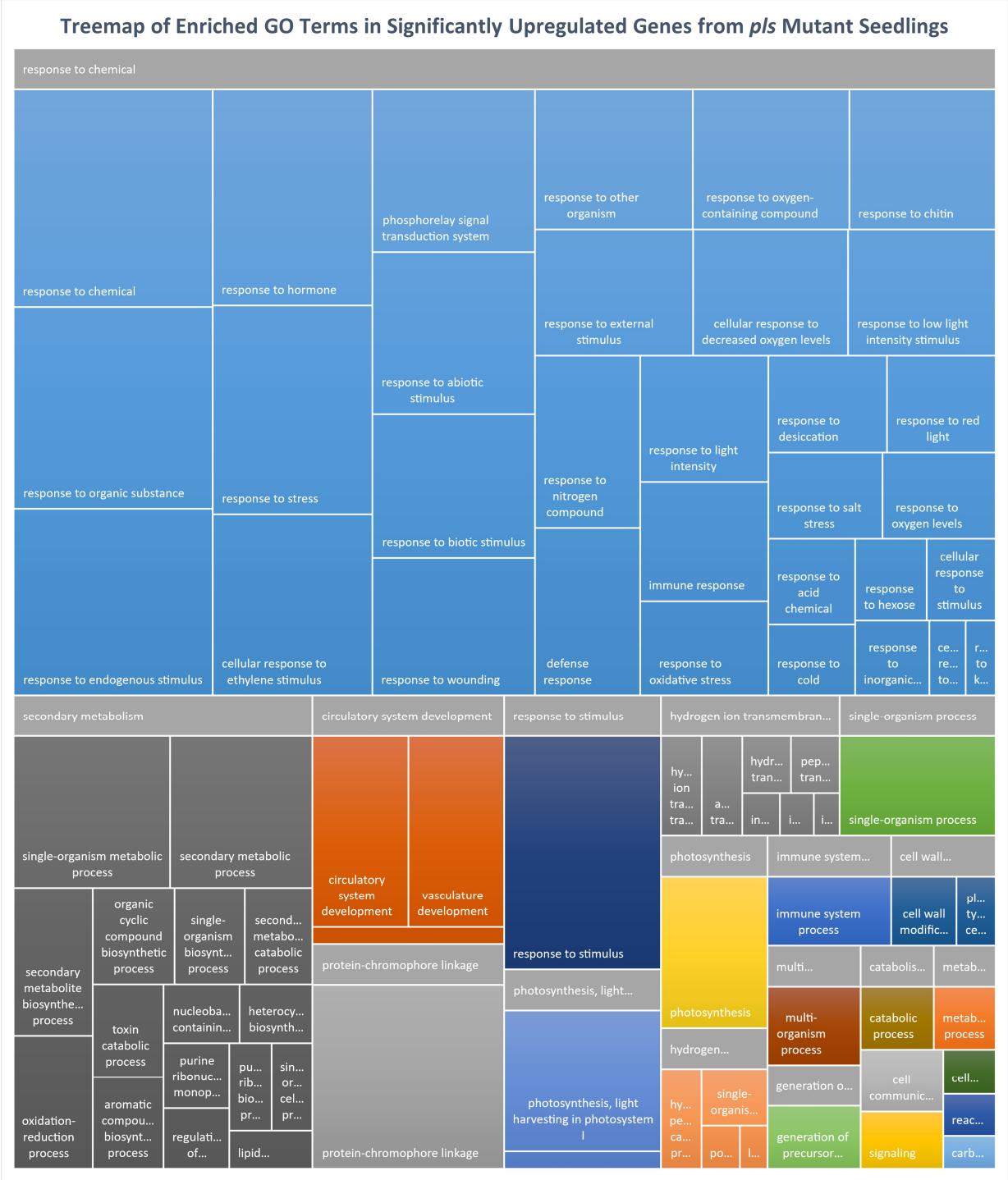


Figure 3.6. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly upregulated genes in *pls* mutant RNA sequencing data, with adjusted P value <0.1 and $\log_2$ fold change ($\log_2$fc) >0.5. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value (|$\log_{10}$ q-value|) of its cluster.

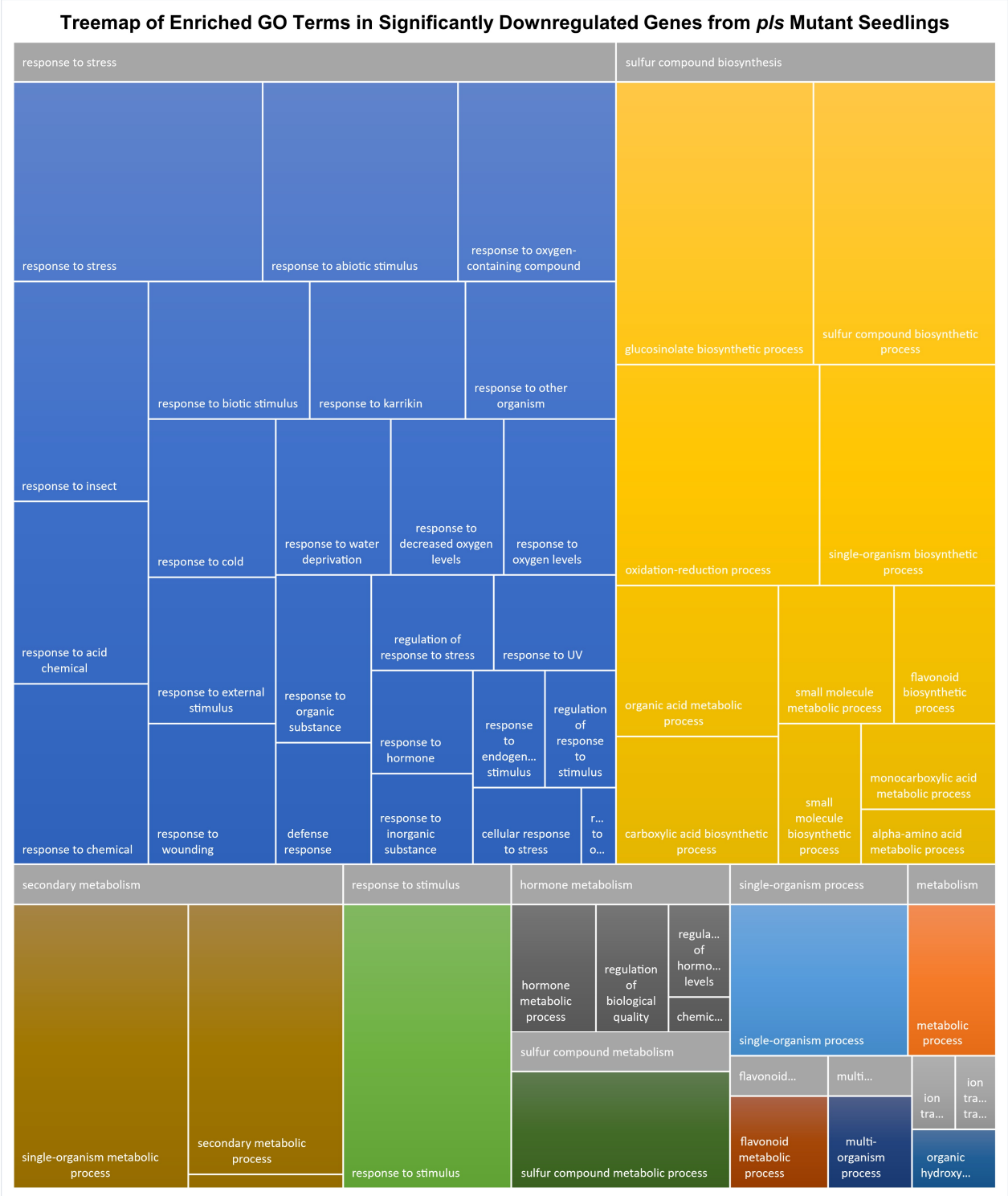


Figure 3.7. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly upregulated genes in *pls* mutant RNA sequencing data, with adjusted P value <0.1 and $\log_2$ fold change ($\log_2$fc) <-0.5. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value ($|\log_{10}$ q-value$|$) of its cluster.

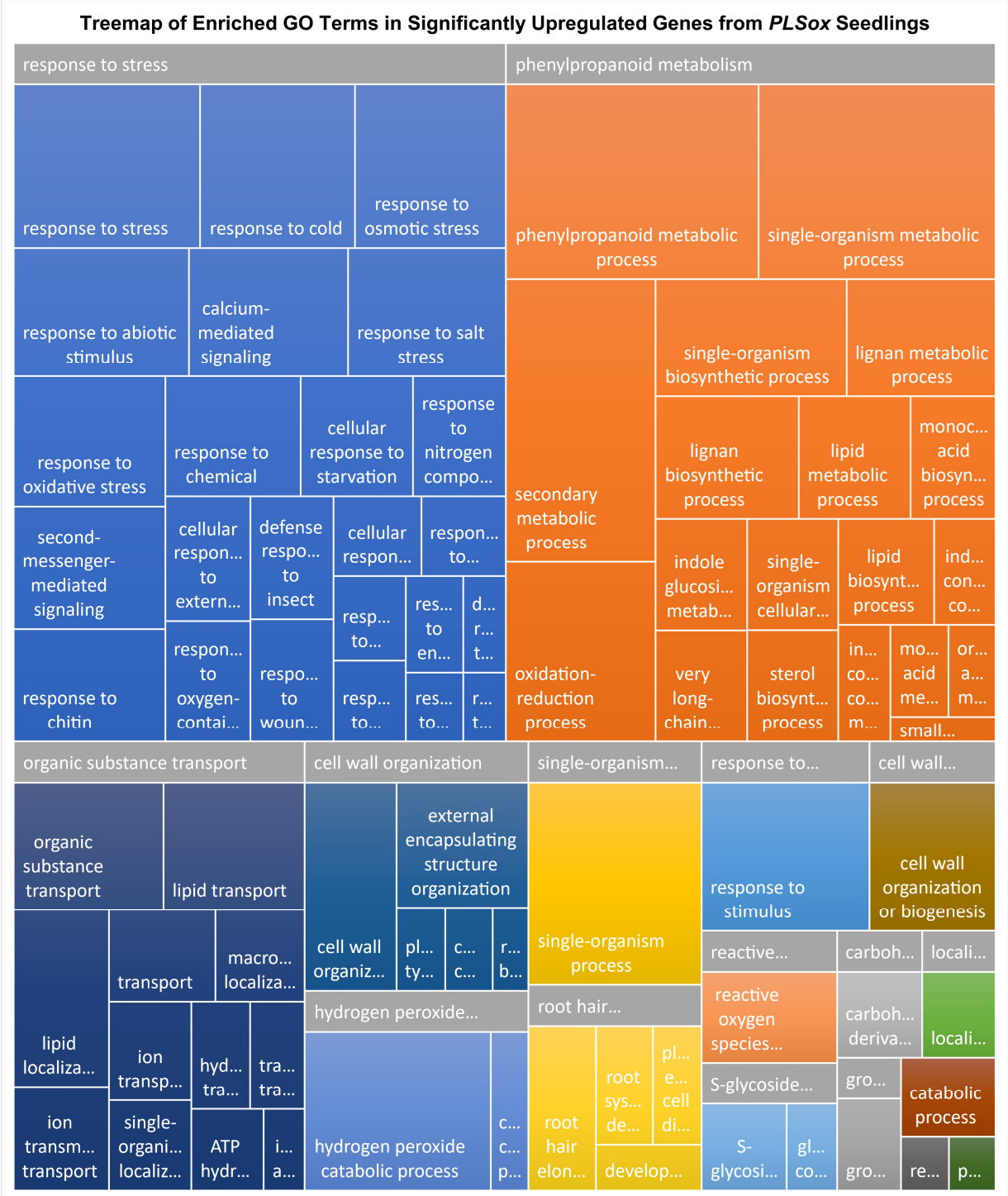


Figure 3.8. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly upregulated genes in PLSox RNA sequencing data, with adjusted P value <0.1 and $\log_2$ fold change ($\log_2$fc) >0.5. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value (|$\log_{10}$ q-value|) of its cluster.

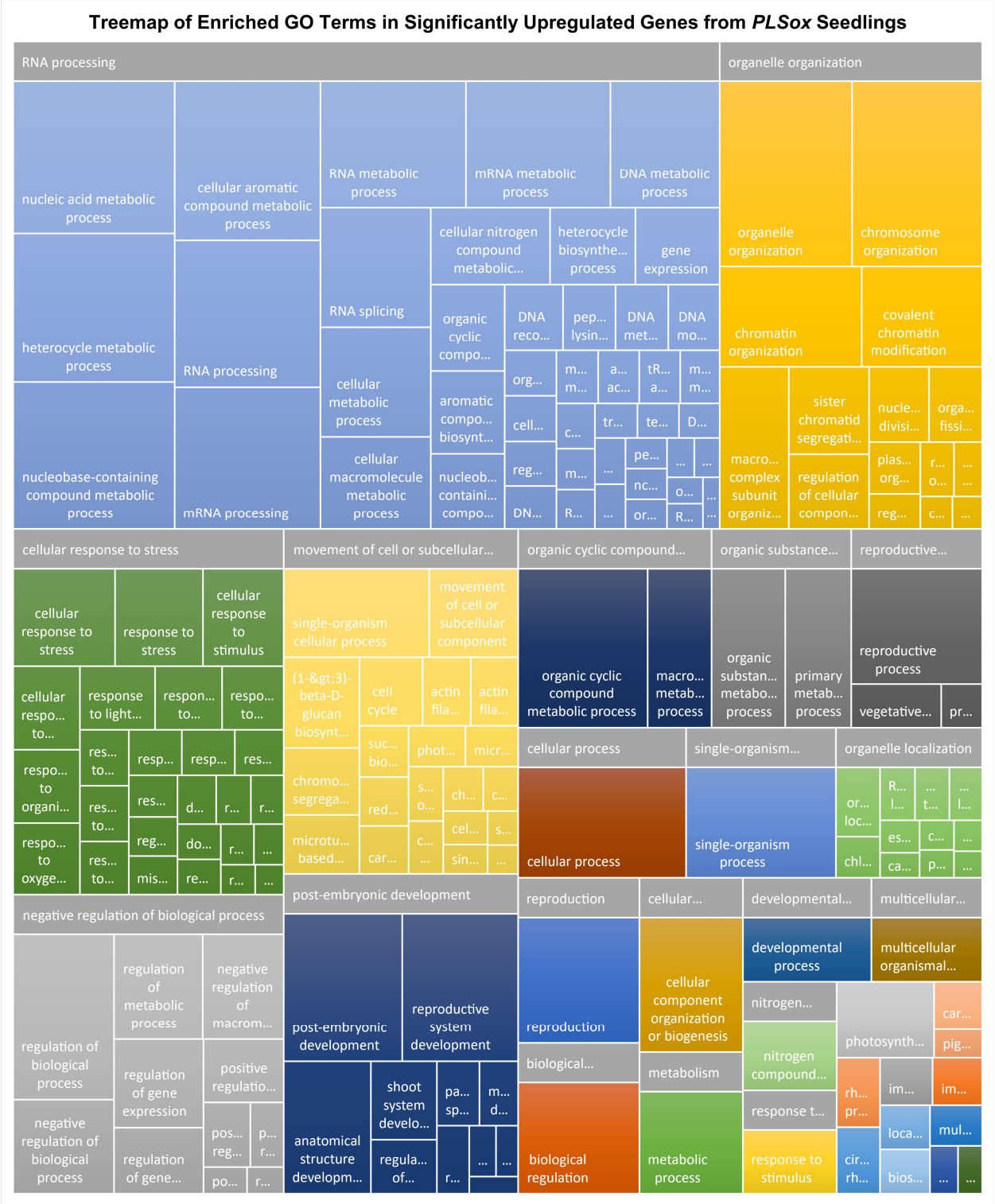


Figure 3.9. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly downregulated genes in PLSox RNA sequencing data, with adjusted P value <0.1 and $\log_2$ fold change ($\log_2$fc) >0.5. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value (|$\log_{10}$ q-value|) of its cluster.

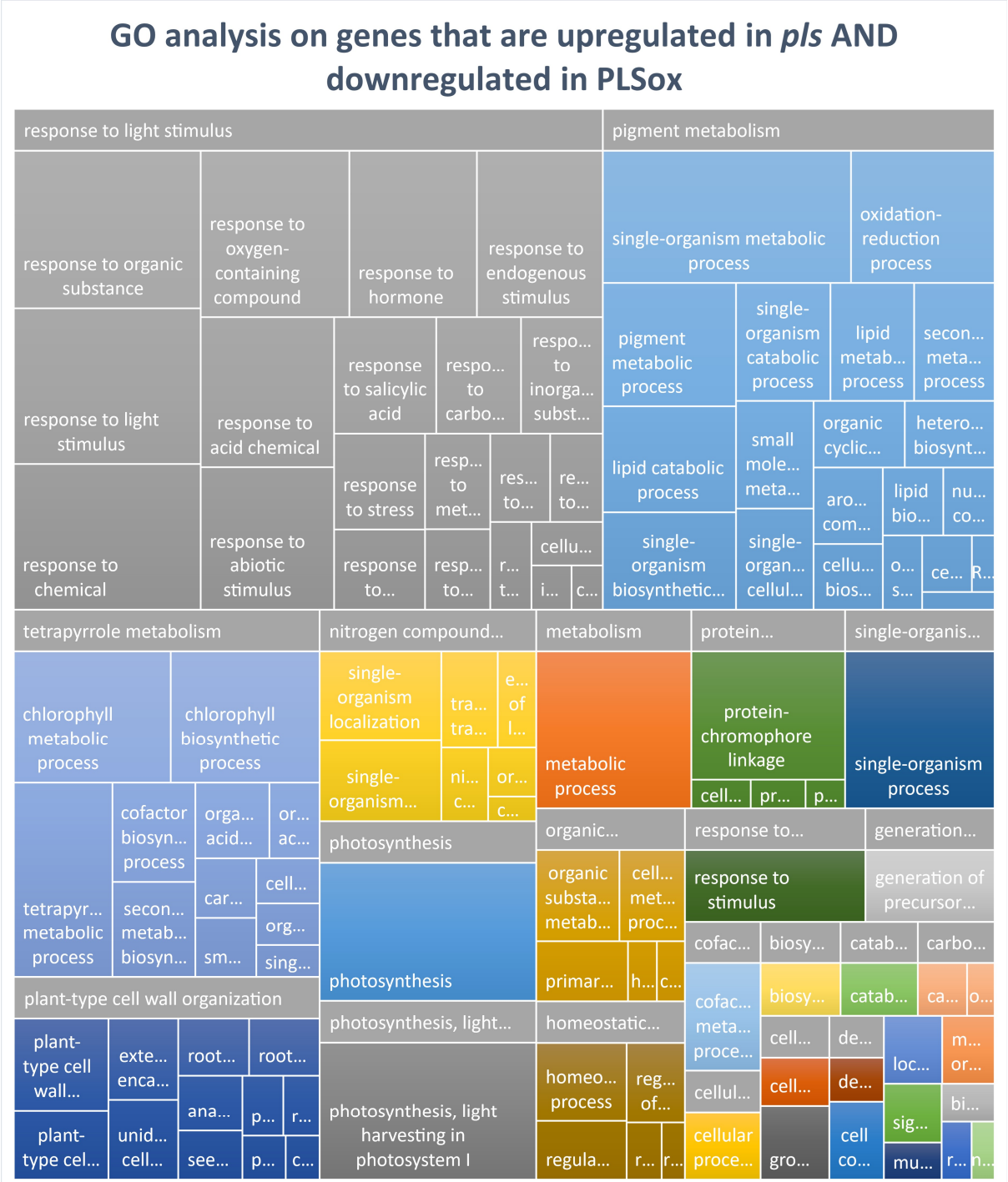


Figure 3.10. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the genes that are significantly upregulated in *pls* **AND** downregulated in PLSox, with adjusted P value <0.1 and $\log_2$ fold change ($\log_2$fc) >0.5. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value ($|\log_{10}$ q-value$|$) of its cluster.

3.3.2 Changes in gene expression in the *pls* mutant

3.3.2.1 Ethylene-related genes are upregulated in the pls mutant

Previous experiment results showed that by inhibiting the enhanced ethylene signalling observed in *pls* mutant seedlings, the short root phenotype was restored to approximately wild-type level. This suggests that increased ethylene signalling is a major contributing factor in the *pls* short root phenotype. This view is further supported by gene ontology (GO) analysis using AgriGO and REVIGO. It is found that among all the significantly upregulated genes in the *pls* mutant, almost all of the top 20 enriched GO terms are related to ethylene-related biological processes, including response to a range of stress-causing stimuli, and is consistent with previous experimental data.

Predictably, the data from the overexpression of *PLS* showed the opposite result. Many stress and defence related GO terms are enriched in the down-regulated gene dataset, while others remain not significantly changed. This finding is in line with the fact that the overexpressor has a similar phenotype to wildtype seedlings under standard growth conditions.

3.3.2.2 No PLS mRNA found in pls transcriptome

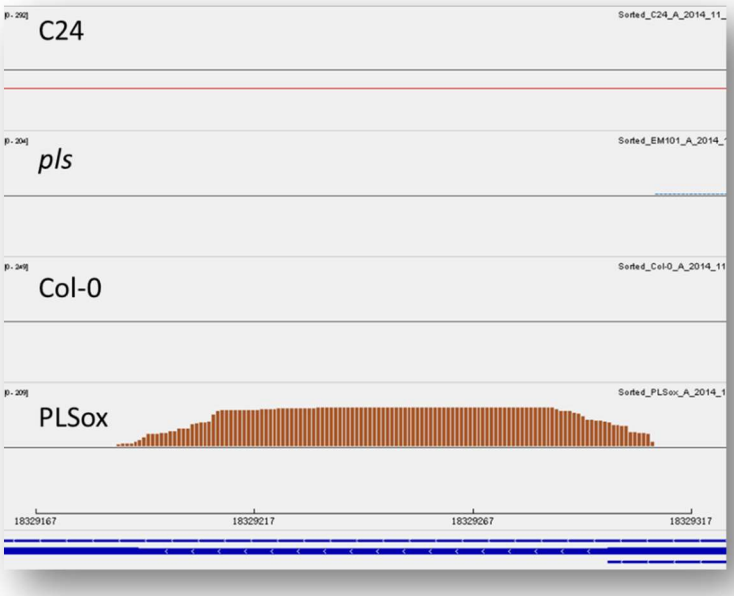


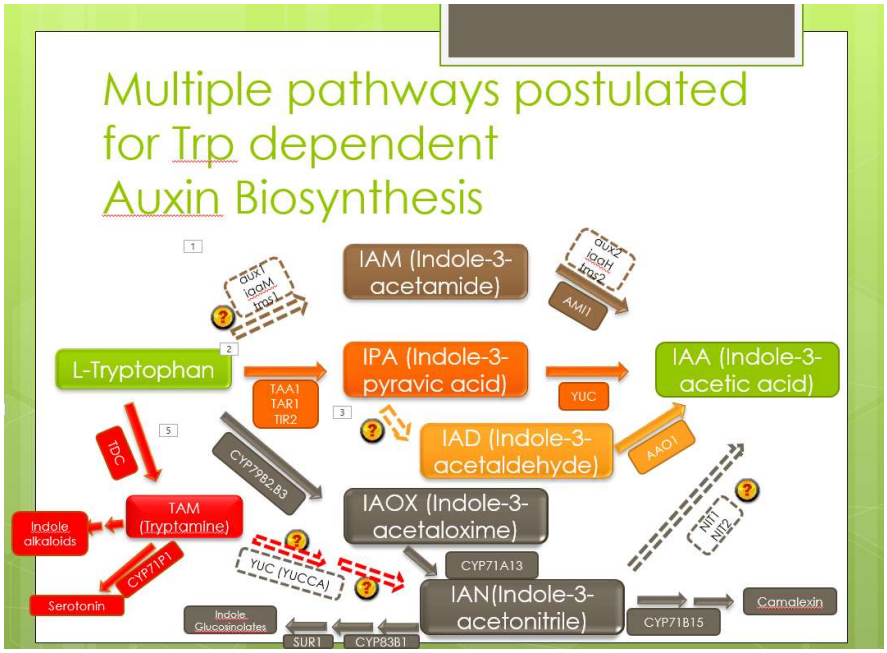
Fig. 3.12 Sashimi Plot showing the transcription level of PLS gene in C24, *pls*, Col-0, and PLSox genotypes. Each bar represents a base in the gene, and the height of the bars represents relative level of expression of each base.

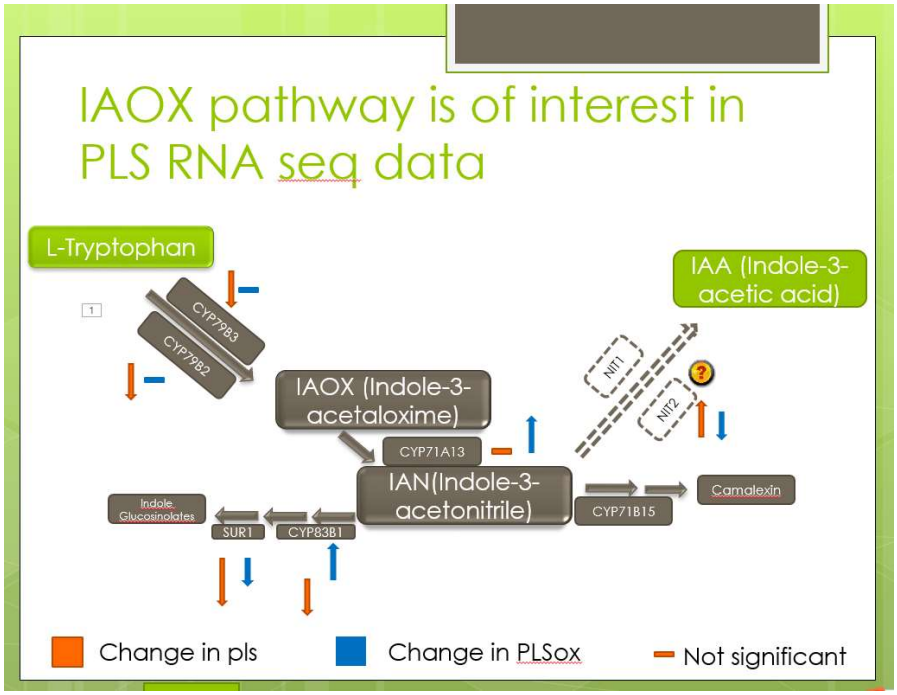
Despite having key roles in signalling pathways, PLS gene has extremely low level of expression in the wildtype transcriptome. Among the samples sequenced, the *PLS* gene had very low levels of expression in both wild types and none in the *pls* mutant, while there is a significant level of *PLS* expression in PLSox lines (Figure 3.11). This confirms that *pls* transgenic line is a knock-out mutant, and that PLSox does produce excessive *PLS* mRNA.

3.3.2.3 Expression of genes in IAOX pathway

In order to study how auxin biosynthesis pathway is affected by POLARIS, relevant genes are looked at in the RNA sequencing data. Among all the pathways postulated in the literature, the IAOX pathway stands out as having most of its genes differentially expressed in *pls* mutant.

A.





B.

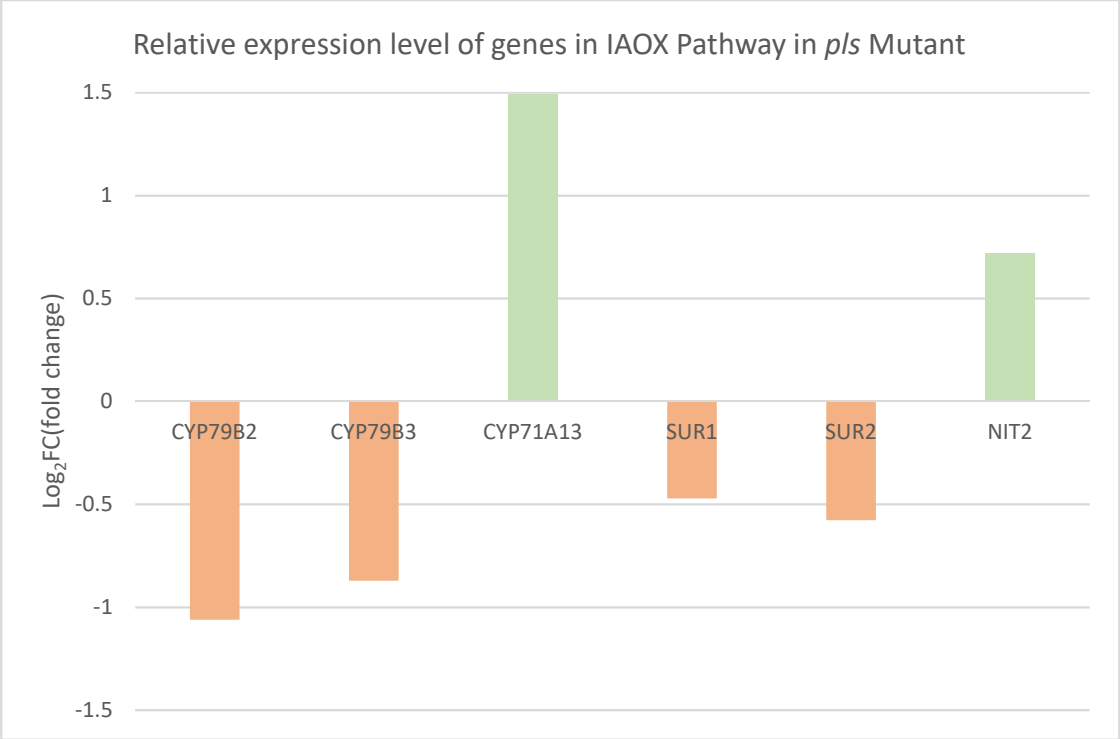


Figure 3.13 A. The change of gene expression level in IAOX auxin biosynthetic pathway. The orange arrows on the genes show the change in the *pls* mutant, and the blue arrows represent the change in the PLSox line. Hyphens represent no significant change in data. **B. The relative expression level of key genes in the IAOX pathway plotted using RNA-seq data from *pls* mutant.** Vertical axis represents log2FC (fold change), where +/- 1 represents two folds up/down regulation relative to that of wild type. q-value < 0.01 for all but SUR1, where q-value < 0.1.

3.3.2.2 PLS has a role in regulating photosynthetic genes

Among all the differentially expressed genes in both *pls* mutant and PLSox, 212 genes were upregulated in *pls* mutant while downregulated in the PLSox. Gene ontology analysis on these genes revealed that the overexpression of *PLS* represses the expression level of many photosynthetic genes, which suggests that PLS might play an inhibitory role on regulating development of photosynthetic apparatus.

This hypothesis is supported by previous studies on the localization of *PLS* expression in young seedlings. Being a regulatory peptide, PLS has extremely low abundance on both transcription and translation level. However, GUS staining experiments clearly revealed that *PLS* transcription is primarily found in non-photosynthetic tissues including root meristem

region and leaf vascular tissue (Casson et al., 2002), consistent with a role for PLS in repressing photosynthetic development or activity.

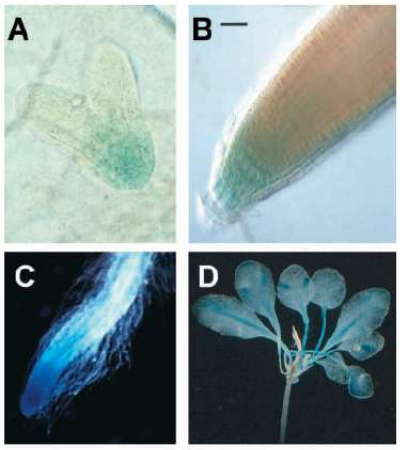


Figure 3.14 GUS expression in the PLS-GUS Promoter Trap Line showing localisation of PLS expression in: (A) heart stage embryo with 5 hr of GUS staining; (B) seven-day-old seedling root tip, with 5 min of Gus staining; (C) seven-day-old seedling root tip with 1 hr of GUS staining; and (D) aerial parts of a 12 day old seedling (Casson et al., 2002).

Much research effort in plant biology is put into looking for ways to increase crop yield by increasing plants' tolerance to various stresses. Studies on drought stress revealed that elevated cytokinin level successfully delayed senescence, leaving the photosynthetic apparatus operational for longer under stress (Smart et al., 1991). Previous experiments showed that the PLS peptide has negative regulatory effects on cytokinin responses and local concentration (Casson et al., 2002, Liu et al., 2010). This suggests that it is possible that PLS's regulatory effects on photosynthesis are achieved at least in part through cytokinin signalling.

3.4 qPCR validation of RNA sequencing data

To test the robustness of RNA sequencing data, qPCR was carried out on additional biological sample replicates. Primers were designed to amplify a selection of genes in the auxin biosynthesis pathways, using cDNA synthesised from RNA isolated from wildtype, *pls*, and *PLSox* tissue samples. This is to validate how the gene expression level changed in *pls* transgenic line and PLS over expressor. All the genes tested showed a similar expression

pattern to those revealed by RNA sequencing data. As an example, two genes in the IAOX auxin biosynthesis pathway were tested again using qPCR. Figure 3.15 shows the relative expression level of these two genes using the data from both qPCR and RNA sequencing experiments, showing similar expression direction of change. Therefore, the qPCR experiment validates the data obtained from the sequencing experiment, confirming its sensitivity and robustness.

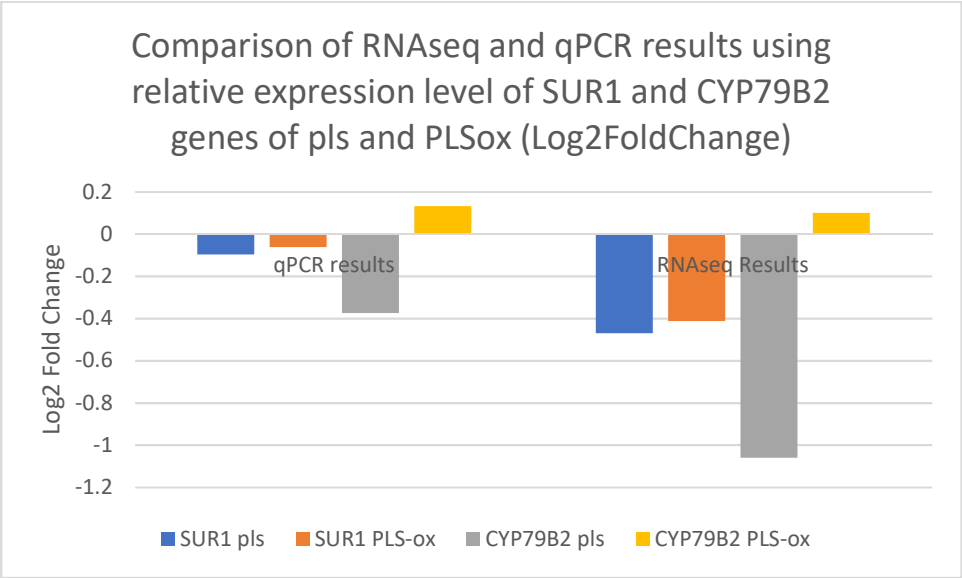
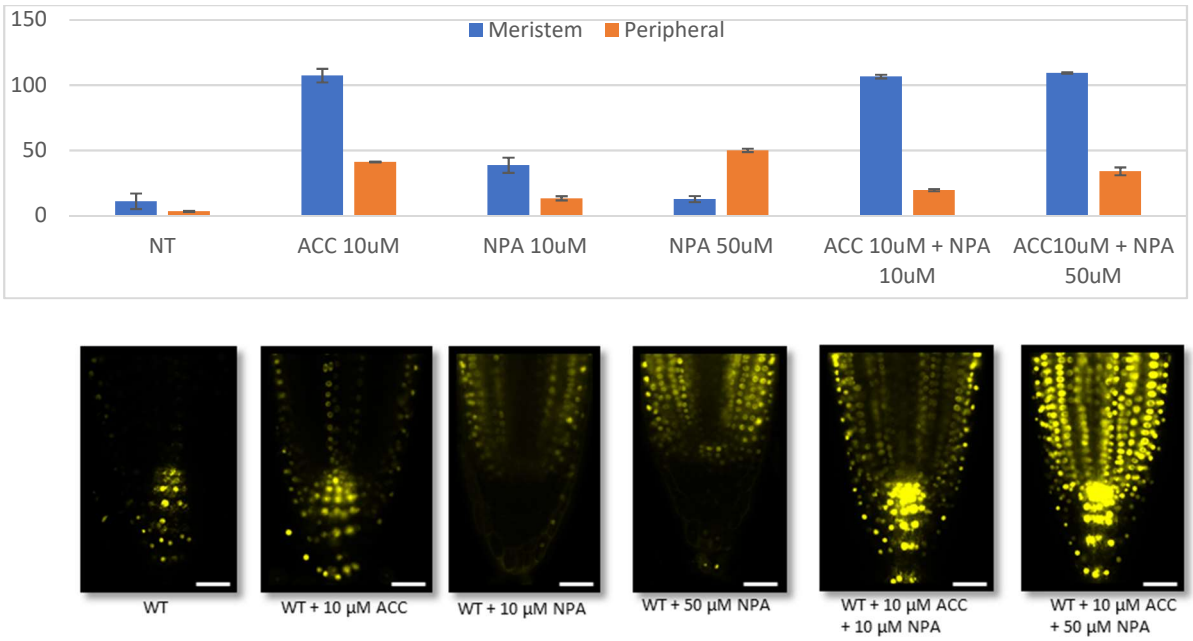


Figure 3.15 Comparison of RNAseq (left) and qPCR (right) results using relative expression level of SUR1 and CYP79B2 genes of *pls* and PLSox. Vertical axis represents the $\log_2$(Foldchange) compared against wildtype expression level. Confidence level $p<0.05$.

3.5 Investigation of a role for PLS in auxin biosynthesis or transport

Previous studies demonstrated that wild type Arabidopsis seedlings treated with exogenous ACC, an ethylene precursor, show a significant accumulation of auxin in root tips. However, the root tip of *pls* transgenic seedlings shows no increase in auxin level under the same treatment (Chilley et al., 2006, Liu et al., 2010). This was confirmed experimentally again in this study (Figures 3.16 -3.19). This observation led us to hypothesize that POLARIS might be an important link between the ethylene signalling pathway and the control of auxin accumulation in the root tip.

A:



B:

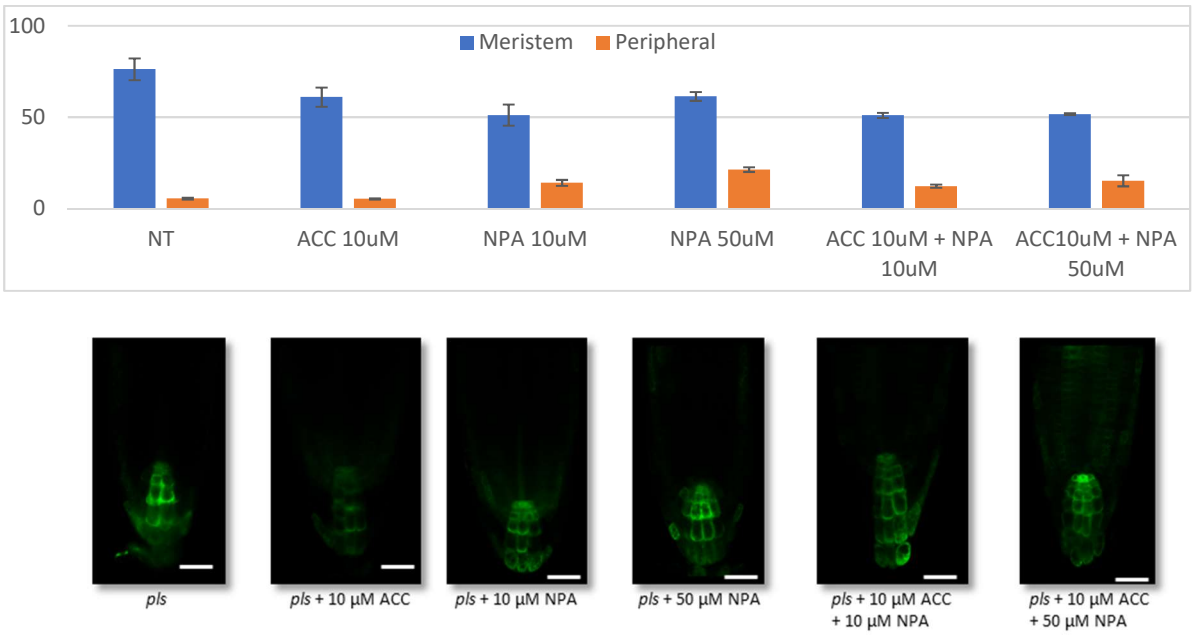


Figure 3.16. Auxin activity in root tips of seedlings grown for 7 days on media with treatment of ACC or NPA. **Block A** pictures are from Col-0 (WT) DR5:Venus, and **Block B** pictures are from *pls* DR5:GFP. The first picture from each panel are non treated (NT). Histograms show measured average (n = 3) intensity level of fluorescent signal in meristem and peripheral region of roots in each graph. Scale bar = 50 µm.

A: **B:**

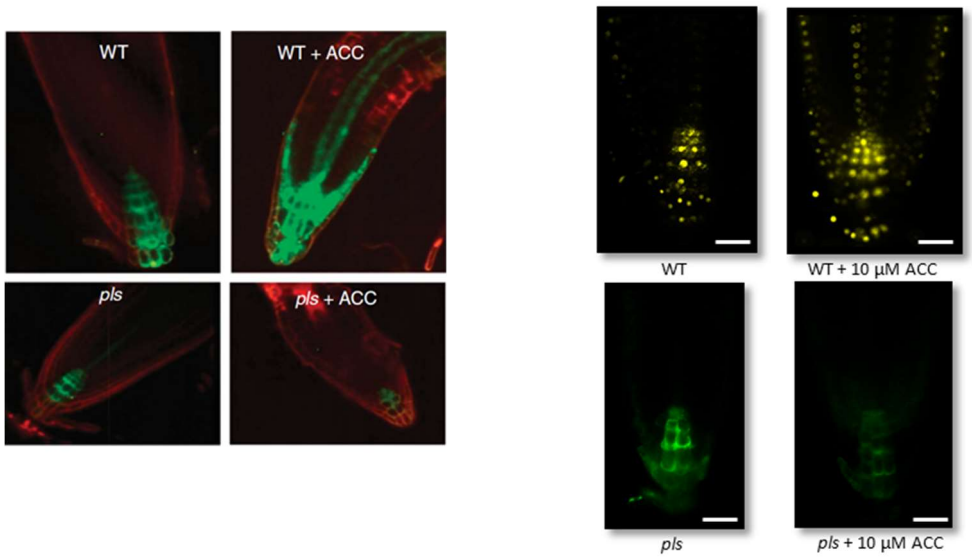


Figure 3.17.

A: Previous experimental demonstration that ethylene precursor ACC induces auxin response in wild-type root tips (revealed as DR5::GFP expression; upper two panels), but the *pls* mutant shows no DR5::GFP induction by ACC (lower two panels). The *pls* mutant (lower left panel) also shows a reduced DR5::GFP signal compared with wild type (upper left panel) (Liu et al., 2010). **B:** ACC treatment on Col-0 DR5::Venus shows elevated auxin level in root tip. ACC treatment on *pls* DR5:GFP. Showing little difference between control and treated Scale bar = 50 µm. Please refer to histograms in Figure 3.16 for fluorescent intensity measurements.

The spatial distribution of auxin in the root tip is maintained by a range of factors including both auxin transport in and out, and local biosynthesis and degradation (Chilley et al., 2006). To investigate the molecular mechanism regulating how auxin accumulation is affected by PLS, it is important to separate the effects on auxin transport and local biosynthesis. NPA is known to inhibit the activity of PIN proteins, which are major components of auxin polar transport mechanism (Forestan and Varotto, 2012, Katekar and Geissler, 1980, Reed et al., 1998). Therefore, NPA is used here to constrict auxin transport. This allows the testing of the hypothesis that, if the low auxin content of the *pls* mutant root tip is due to enhanced auxin transport out of the root (as the PINs are enhanced in the *pls* mutant; Liu et al. 2013), then inhibition of transport (by NPA) should lead to the accumulation of auxin in the *pls* tip, assuming PLS is not required for auxin biosynthesis. If PLS is required for auxin biosynthesis,

we would expect to see no increase in auxin in the root tip in response to ACC, in the presence of NPA.

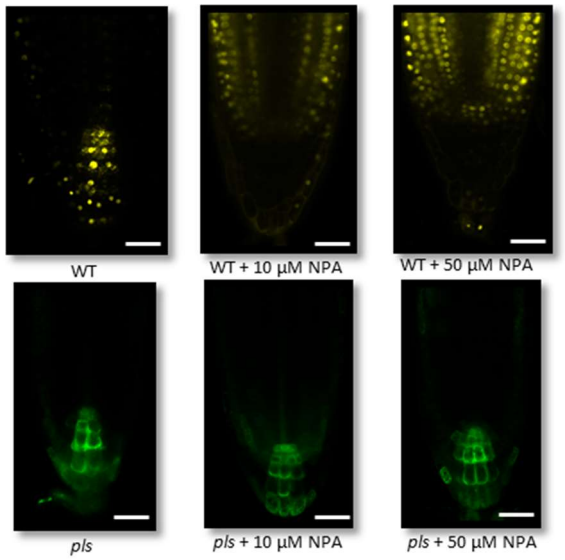


Figure 3.18. UP: Auxin distribution in root tip is dramatically changed in WT seedlings treated with NPA. **DOWN:** Effect of EPA treatment on auxin distribution of pls mutant seedlings. Scale bar = 50 µm. Please refer to histograms in Figure 3.16 for fluorescent intensity measurements.

Wildtype seedlings treated with NPA show dramatic change in auxin distribution in the root tip. As shown in Figures 3.17 and 3.18, the DR5::Venus fluorescence (yellow areas in upper panels), representing auxin response, shifted its distribution from mainly in the columella and quiescent centre to cortex and epidermis regions, with a dosage dependent effect. In the *pls* mutant, NPA treatment showed a much less significant change in auxin distribution (green areas in lower panels) in the root tip. In contrast to the change seen in the wild type, the quiescent centre and columella region of *pls* retained its auxin signal, while the cortex and epidermal region gained a slight but significant elevation. This suggests that in wild type seedlings, a significant proportion of active auxin in the quiescent centre and columella region in root tip is transported up to the upper cortex and epidermal tissues. By treating seedlings with NPA, transport of auxin is restricted, leading to accumulation of auxin in the peripheral region; and less auxin concentration in the quiescent centre and columella region.

When ACC and NPA are applied to seedlings simultaneously, both wild type and *pls* mutant seedlings showed increased auxin activity in root tip, with wild type having more severe effect. This shows that ACC induces local auxin biosynthesis both in the meristem region and peripheral region in wild type seedling roots. However, the lack of significant elevation of auxin accumulation in *pls* seedlings suggests that ACC failed to induce local auxin biosynthesis in the root tip, and that PLS peptide is required for ethylene-mediated auxin biosynthesis.

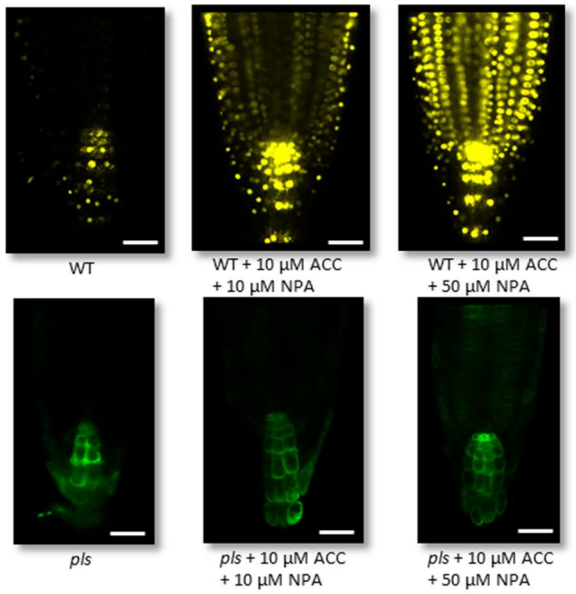


Figure 3.19. UP: Auxin distribution in root tip is dramatically changed in WT seedlings treated with NPA and ACC simultaneously. **DOWN:** Effect of NPA and ACC joint treatment on pls mutant seedlings). Scale bar = 50 µm. Please refer to histograms in Figure 3.16 for fluorescence intensity measurements.

Summary

The aim of this chapter was to investigate the function of POLARIS and its relationship with other hormones and genes, and its role in auxin biosynthesis pathways. Using RNA sequencing technology, the whole transcriptomes of Arabidopsis variants including WT, *pls* and *PLSox* were obtained to look at the relationships at the gene expression level. Gene ontology analysis was carried out to provide a general overview of the large amount of data generated in RNA-seq. Other bioinformatic tools, bioimaging, literature search, and other analyses were carried out to further investigate the data. As the aim of this chapter is particularly aimed at the role of POLARIS in auxin biosynthesis, much work has been focused on identifying its effect independent of ethylene-induced auxin responses.

The data from RNA-seq experiments revealed that PLS peptide is highly likely to be involved in the ethylene response as the lack of PLS in *pls* mutant seedlings induced upregulation of a significant group of ethylene related genes. The analysis also suggests that PLS may be playing a role in regulating development of the photosynthetic apparatus, at least in part through interaction with cytokinin. The bioimaging experiments confirmed that PLS peptide is required for ethylene-mediated auxin biosynthesis in the root tip.

Chapter 4. Function and Relationships of MDF

4.1 Introduction

MERISTEM-DEFECTIVE (*MDF*) is a gene which was previously identified in an embryogenesis analysis using laser capture microdissection technology (Casson et al. 2009). The gene encodes a predicted polypeptide of 820 amino acids, forming a putative serine-arginine related argine-serine (RS) domain protein. It was found to be crucial for correct root meristem organisation and maintenance. Two independent homozygous mutants (*mdf-1* and *mdf-2*) of MDF exhibit three cotyledons in seedlings, in addition to some similar phenotypes to *pls* seedlings, including delayed development and short roots. Most interestingly, the mutant seedlings also have a reduction in auxin levels in the meristem region and disturbed auxin patterning in the root tip region, associated with impaired development. Based on its structural homology to the human SART1 and yeast snu66 proteins (Makarova et al. 2001), it is hypothesised that the MDF protein may plays a key role in regulating RNA splicing, which is a fundamental mechanism that expands diversity in the transcriptome and proteome; correctly spliced genes are basic requirements for normal cell activity.

4.2 Next-Gen RNA sequencing on MDF

The goal of this RNA sequencing experiment was to identify the changes in expression levels, as well as alternative splicing patterns, of RNAs as a result of mutating the *MDF* gene in Arabidopsis. To understand how the mutation has affected RNA profiles linked to altered root development, especially in the meristem region, samples were taken from 7 day old seedlings.

Sequencing was performed on samples from Col-0 as wildtype control, both mutants *mdf-1*and *mdf-2*, and a transgenic over-expresser MDFox. Each genotype was sequenced as three independent biological replicates, resulting in a total of 12 samples. Two lanes of high output run on Illumina HiSeq 2500 platform were performed to achieve up to 400-500 million paired-ends reads per lane, equivalent to 65-78 million paired-ends reads per sample. This level of read depth was expected to be enough to analyse differences in alternative splicing events between samples. A prediction is that the *mdf* mutants would show not only differentially expressed genes associated with meristem function, but also

evidence of mis-splicing of at least some RNAs, given the predicted function of MDF as a spliceosome factor.

4.2.1 Quality control.

The program FastQC was used to test the quality of sequencing data. The data obtained from this experiment achieved exceptional quality. According to FastQC reports, the quality score of most reads lie on top of the quality scale, which is especially useful for detecting small RNA molecules and alternative splicing events.

4.2.2 Analysis of differentially expressed genes using RNA-Seq Data

The experiment has generated high quality and depth data, which in turn gives a strong foundation for identifying differentially expressed genes. The differentially expressed genes (DEGs) are selected based on an adjusted p-value (q-value) <0.05, representing the confidence level; with a log2fc >1 or <-1 representing more than two fold change in up or down direction respectively.

The results shows that the severely dwarfed and deformed phenotypes of *mdf-1* and *mdf-2* are associated with significantly altered expression levels of thousands of genes. Compared with wild type Col-0, there were 4195 up-regulated genes and 5404 down-regulated genes in *mdf-1*. For *mdf-2*, which has a less severe phenotype than *mdf-1*, the number of up and down regulated genes are 2830 and 3449. Both of the figures in *mdf-2* are considerably less than that of *mdf-1*, which is in line with the stronger phenotypes observed in *mdf-1* seedlings, suggesting that the truncated MDF protein in *mdf-2* might still be partially functional.

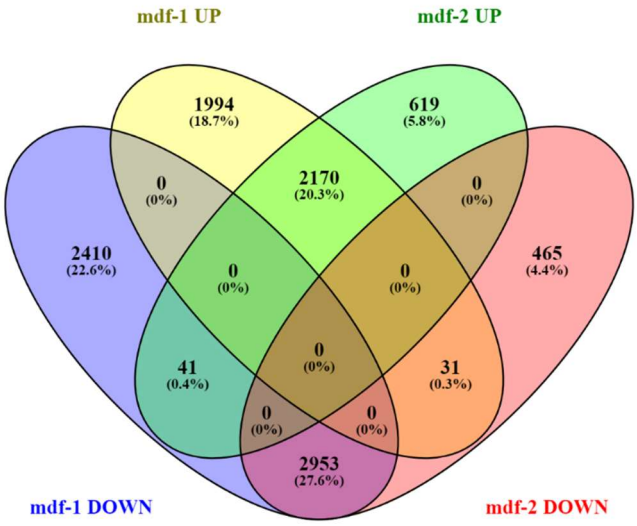


Figure 4.1 Venn diagram showing the number of differentially expressed genes (DEGs) in *mdf-1* and *mdf-2* homozygous transgenic 7 d.a.g. seedlings following the RNA sequencing experiment, with adjusted p value < 0.05, and $\log_2$ fold change ($\log_2$fc) >1. Each oval contains all up or down regulated genes in one of the genotypes, and the overlapping parts represent numbers of genes meeting the conditions of more than one encircling oval. Percentage under each number is calculated by dividing each number by the total number of DEGs in the diagram.

4.2.3 Overview of Differentially Expressed Genes using Gene Ontology (GO) Analysis

High-throughput experiments can produce large amounts of data which are difficult to comprehend. As a good example, the RNA sequencing experiment in this project identified thousands of differentially expressed genes from each sample. In order to interpret such amount of data, gene ontology (GO) enrichment analysis was carried out using the online platform agriGO. This analysis identifies frequently occurring GO terms within the list of DEGs to show the biological processes that are affected in by the mutation in *mdf-1* and *mdf-2* seedlings. However, the list of enriched GO terms can still be long and redundant, making it hard to interpret. To make the result more intelligible, the long list of enriched GO terms is further reduced using REVIGO, which groups GO terms semantically based on their functional similarity into clusters, using a single, most relevant GO term picked from each cluster as the representative for that cluster. To visualise the output, the clusters are displayed in a tree map as rectangles, which are grouped further into superclusters each assigned a colour, with the sizes of the rectangles represent the p-values of the enriched GO terms. This analysis gives an overview of the DEG profile of each sample, making it easier to identify the potential biological processes and pathways of interest.

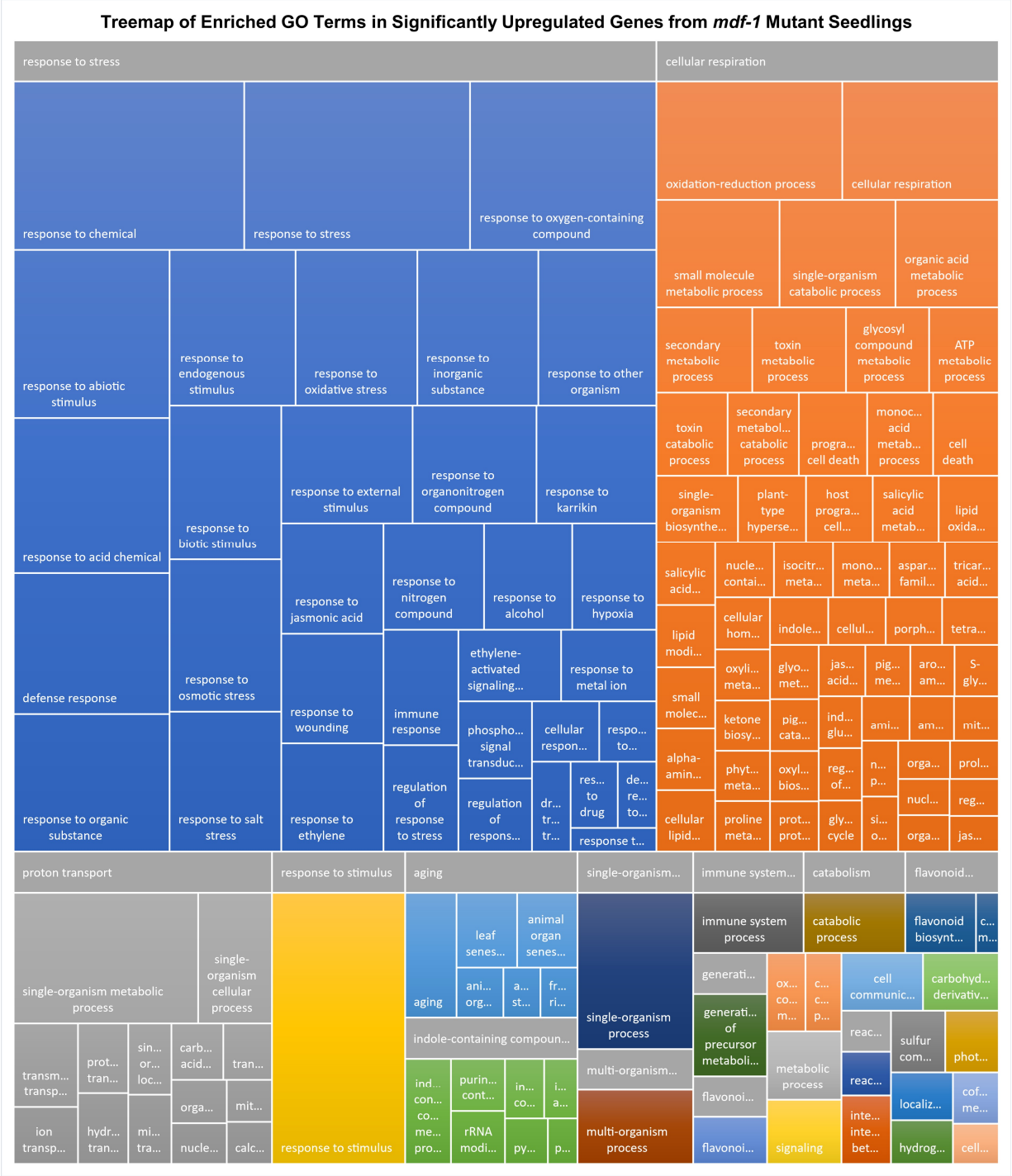


Figure 4.2 Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly upregulated genes in *mdf-1* mutant RNA sequencing data, with adjusted P value <0.05 and $\log_2$ fold change ($\log_2$fc) >1. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value ($|\log_{10}$ q-value$|$) of its cluster.

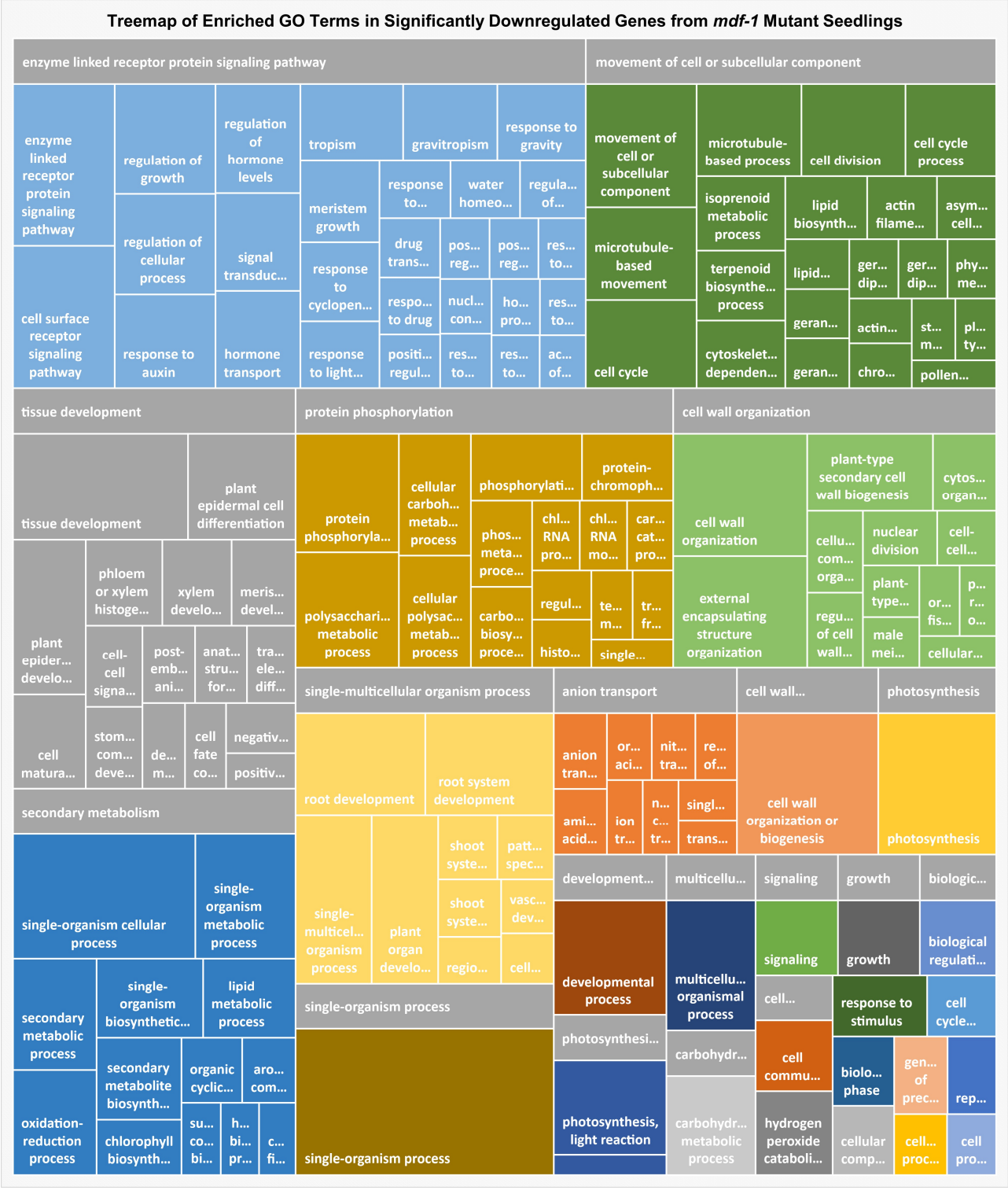


Figure 4.3 Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly downregulated genes in *mdf-1* mutant RNA sequencing data, with adjusted P value <0.05 and $\log_2$ fold change ($\log_2$fc) <1. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value (|$\log_{10}$ q-value|) of its cluster.

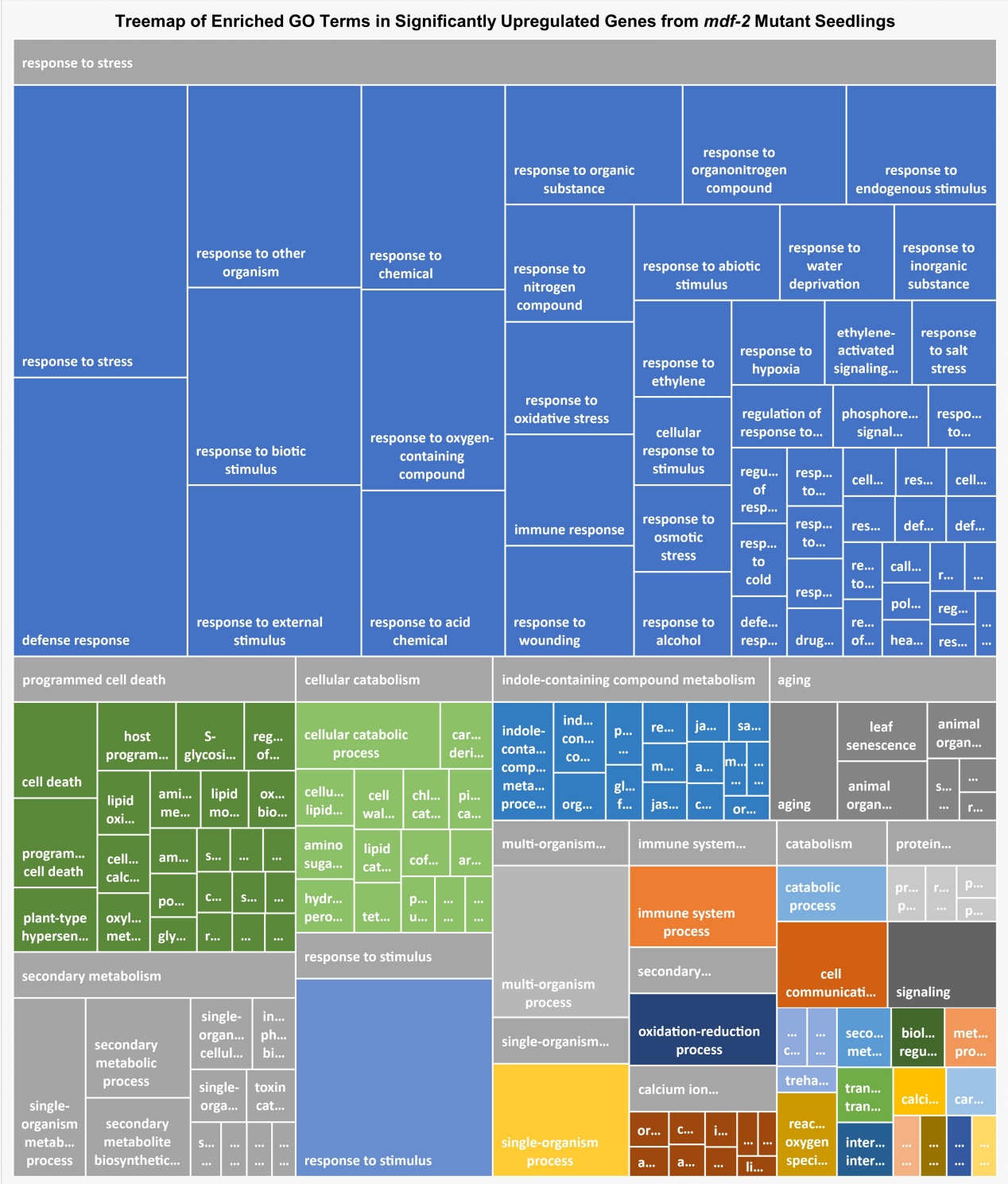


Figure 4.4 Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly upregulated genes in *mdf-2* mutant RNA sequencing data, with adjusted P value <0.05 and $\log_2$ fold change ($\log_2$fc) >1. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value (|$\log_{10}$ q-value|) of its cluster.

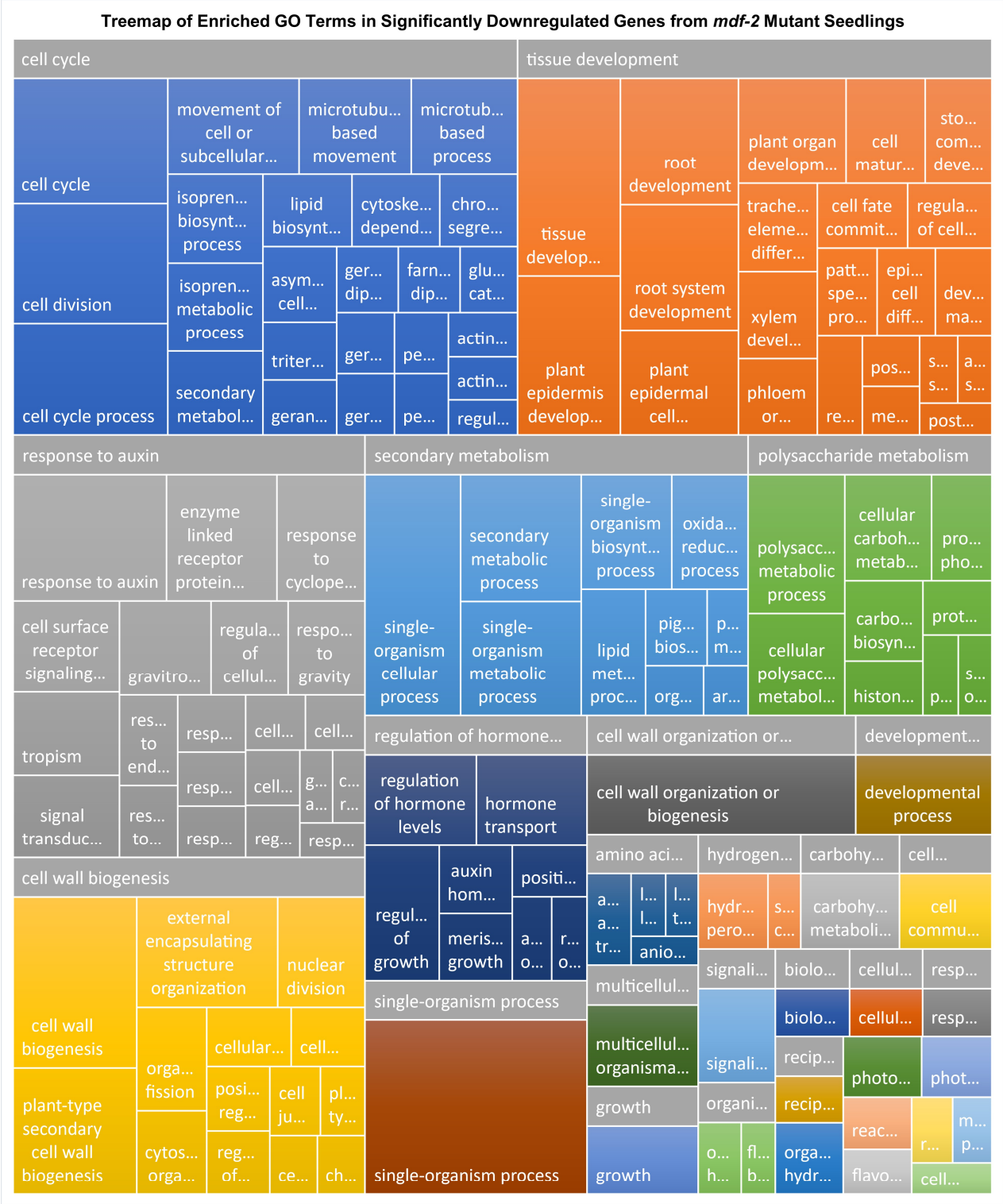


Figure 4.5 Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the significantly downregulated genes in *mdf-2* mutant RNA sequencing data, with adjusted P value <0.05 and $\log_2$ fold change ($\log_2$fc) <1. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value ($|\log_{10}$ q-value$|$) of its cluster.

4.2.*4* Key findings from the gene ontology analysis

The tree maps show a strong pattern in both *mdf-1* and *mdf-2* transcriptomes that both exhibit upregulation in stress related genes, as seen in Figures 4.4 and 4.5, and the enriched GO terms include response to stress, defence response, response to other organism, immune response, aging, and reactive oxygen species metabolism (ROS). As one of the key stress related pathways, response to ethylene also shows up in both upregulated GO enrichment lists. This suggests that mutation of *MDF* genes has profound effects on many biological processes in *Arabidopsis*, and many stress responsive pathways were activated or upregulated in an effort to compensate for the damage. In both tree maps, aging and programmed cell death were also shown significantly upregulated, which might be a result of various stress placed upon the young seedlings.

The supercluster 'response to stress', which also contain 'response to ethylene', is the largest one in the list in both up regulated tree maps. This suggests that MDF might be playing a negative regulatory role in stress response. Major superclusters also contain 'programmed cell death', 'secondary metabolism', and 'aging'. The supercluster 'indole-containing compound metabolism' suggests that auxin biosynthesis might be affected in mutant lines.

In both downregulation tree maps for *mdf-1* and *mdf-2* samples, most of the enriched GO terms are related to basic and crucial cellular functions and developmental processes. Both figures show down regulation in genes in GO terms of enzyme linked receptor protein signalling pathway, which is one of the most fundamental pathways which transmit signals across membrane by transmembrane receptors binding extracellular ligands causing enzymatic changes in the cytoplasm (Alberts B, 2002). One of the important pathways using this principal is the auxin signalling pathway, which is also enriched in both downregulation tree maps under GO term response to auxin. Other downregulated GO terms include protein phosphorylation, photosynthesis, cell wall organisation, tissue development and growth.

The GO enrichment analysis provides a general overview of the change in transcriptome and gives indication on which direction should be investigated in more detail. These findings are in line with the short root and delayed development phenotypes of *mdf-1*and *mdf-2* homozygous mutants, as well as previous experiment results, such as the low expression level of auxin transporter PIN protein genes.

4.2.4.1 PIN protein Genes are supressed in mdf-1 seedlings

The PIN protein family is a group of auxin carriers, which are crucial for maintaining the correct auxin distribution in the root for its development and growth. The quiescent centre in the stem cell region of the meristem is responsible for maintaining the stem cell population that surrounds it, and which requires a peak of auxin concentration.

Gene ID	Name	baseMean	mdf-1 log2 FoldChange	q value	mdf-2 log2 FoldChange	q value
AT1G73590	PIN1	425.33592	-2.10	3.70E-38	-1.38	4.17788E-18
AT5G57090	PIN2	763.87419	-2.81	2.43E-49	-1.10	4.10648E-09
AT1G70940	PIN3	1880.1384	-1.44	8.81E-34	-1.55	2.03376E-38
AT2G01420	PIN4	1280.8696	-2.14	9.87E-42	-1.82	2.39411E-30
AT5G16530	PIN5	33.033236	-3.47	5.44E-13	-1.82	2.17362E-06
AT1G77110	PIN6	46.301212	-3.40	1.75E-13	-2.14	4.10881E-07
AT1G23080	PIN7	3048.8329	-2.69	4.79E-150	-1.28	1.66147E-36
AT5G15100	PIN8	0.7652549	-0.19	8.64E-01	0.61	NA

Figure 4.6 Table showing the relative expression level of PIN genes in *mdf-1/2* mutant using data from RNA sequencing experiments. Green ticks indicate confidence level greater than 95%.

PIN proteins can be classified into two subfamilies based on similarity in their structures and subcellular localisation. PIN1-4 and PIN7 proteins are members of the larger subfamily, the long PINs, and are characterised by a highly conserved long sequence of hydrophilic loop, connecting two transmembrane domains each containing about 5 hydrophobic regions. They are localised at the plasma membrane and are defined as auxin efflux carriers as they move auxin molecules from the cytoplasm to intercellular space. Although having shortened hydrophilic loop, PIN6 can also be put into this group due to the highly similar transmembrane domains it shares with the rest of the family. But its structure anchors the protein in the ER membrane. The other major subfamily includes PIN5 and PIN8, which are characterised by almost absent hydrophilic loop, and they are also found to be localised in the ER like PIN6 (Krecek et al., 2009).

In the RNA sequencing differential expression analysis, it was found that all PIN protein genes apart from PIN8 are significantly down regulated by 1.5 – 3.5 fold with more than 99.99% confidence level, consistent with results from previous experiment where auxin distribution is found dramatically disrupted in *mdf-1* seedlings (Casson et al., 2009)

4.2.4.2 key genes in meristem development and stem cell activity are down regulated

It is known that MDF is required for the correct meristem patterning in Arabidopsis, one of the most novel phenotypes of *mdf* mutant seedlings is the absence of organised meristem region. With correct auxin concentration distribution, specifically a concentration maximum at the root tip region, the *PLETHORA* family genes are expressed to establish the quiescent centre, where a group of cells that are surrounded by stems cells that divides and differentiate into necessary cell types in the roots. This process also requires the actions of the SCARECROW (SCR) AND SHORTROOT (SHR) genes which are responsible for establishing radial patterning (Helariutta et al., 2000).

Four of the PLT family members, PLT 1-4, partly sharing their transcriptional domain, function in the root meristem to maintain cell division and repress cell differentiation (Galinha et al., 2007). In this RNA sequencing experiment, all but *PLT3* were found significantly down regulated by about 3-5 fold in both *mdf-1* and *mdf-2* mutant, with *PLT3* down at 0.8 fold. Similarly, *SHR* is also found to be down regulated by about 2.5 fold.

In the *PLT* family, *PLT3, PLT5,* and *PLT7* are known to be active in cells that are developing into lateral roots and in the shoot apical meristem, where they function to control the generation of lateral roots (Galinha et al., 2007). In contrast to *PLT1-4*, the RNA experiments show that both *PLT5* and *PLT7* are significantly up regulated in *mdf1* and *mdf2* mutant seedlings.

Key genes in stem cell regulation that are mis-regulated also include *WOX (WUSCHEL RELATED HOMEOBOX)* gene family. These genes contain homeobox sequence which encode the DNA binding domain, and they are key components in the regulation of embryonic pattern by controlling cell differentiation on timing and location (Haecker et al., 2004). In the RNA sequencing result, *WOX1*, *WOX4*, and *WOX5* are significantly down regulated ranging from 1.5 to 4.1 fold, with *WOX2* showed 1 fold up regulation, while, *WOX3* showed no difference.

4.2.4.3 Key genes in stress response are up regulated

In the tree maps for up regulation in both *mdf-1* and *mdf-2* transgenic lines, the largest supercluster is 'response to stress', which also contains 'response to ethylene'. This suggests that MDF might be acting like a switch, negatively regulating gene expressions in response to stresses. The result also shows that the expression of PLS is significantly down regulated in *mdf-1 and 2*. This is consistent with previous experiments confirming that, being a key component of an ethylene receptor, PLS negatively controls ethylene response.

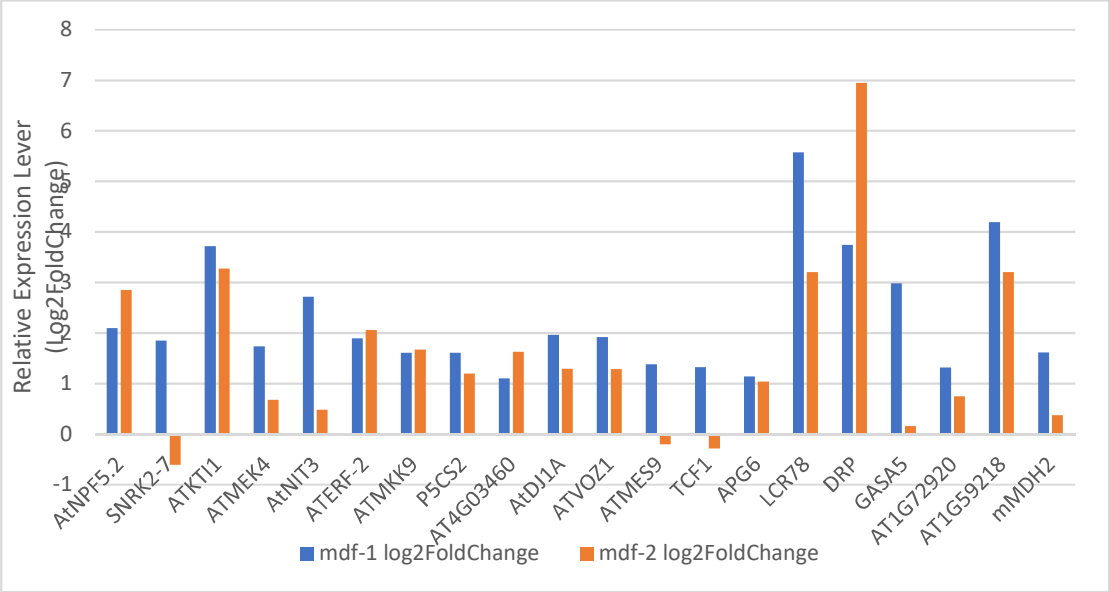


Figure 4.7 Key genes included in gene ontology term 'response to stress' are up regulated in both *mdf-1/2* mutants. (q<0.05).

4.2.5 Alternative Splicing analysis.

Alternative splicing is an essential mechanism where different mRNAs can be generated from a single gene, thus increasing diversity in transcriptome and proteome. Statistics show that more than 95% of human genes undergo alternative splicing at various stage of development (Nilsen and Graveley, 2010). In *Arabidopsis thaliana*, a recent study shows that 82190 unique transcripts were generated from 34212 genes. These transcripts were generated from RNA sequencing data, and assembled into a Reference Transcript Dataset for Arabidopsis, AtRTD2 (Zhang et al., 2017).

MDF is an ortholog of human SART1 and yeast snu66 proteins, which are key components of spliceosomes responsible for RNA splicing. As a result, we hypothesize that the traumatic

phenotype of *mdf-1* mutant seedlings are caused by defects in RNA splicing, and MDF plays a regulatory role in controlling alternative splicing in *Arabidopsis*.

Alternative splicing events are classified into 5 categories. Firstly, Alternative 3' splice site (A3SS) and A5SS are events that have exon spliced at different 3' or 5' end respectively. Then, mutually exclusive exons (MEX) refers to situations where one or the other of two exon is retained, but not both within the same mRNA. Intron retention (IR) is when an intron is retained as part of a transcript, while skipped exons (SK) means that when an exon is simply skipped (Black, 2003).

Alternative splicing analysis was performed using AtRTD2 (Prof. John Brown, University of Dundee) as the reference transcriptome, a total of 2413 splicing events were found to be significantly different in *mdf-1* transgenic lines compared with wildtype. The table below shows the number of differentially spliced events under each of the 5 alternative splicing categories.

AS Type	No. of differential splicing events in *mdf-1*
Alternative 3' splice site (A3SS)	404
Alternative 5' splice site (A5SS)	444
Mutually exclusive exons (MEX)	11
intron retention (IR)	1321
Skipped exons (SK)	233
Total	2413

Figure 4.8 The number of differential splicing events of *mdf-1* transgenic line compared with wildtype, with FDR (false detection rate) < 0.01; ILD (Inclusion Level Difference) > 10% and < -10%.

GO analysis on differentially spliced genes.

The 2413 differential splicing events are found from 1080 unique genes from the RNA sequencing data. Gene ontology analysis is used again to learn the overall pattern among these genes. AgriGO was used to extract GO terms from these genes, before then GO terms were passed onto REVIGO to group into clusters and superclusters and shown in Figure 4.9 as a tree map.

The tree map shows that a large proportion of differentially spliced genes is involved in RNA processing, having 'mRNA metabolism' as the largest supercluster containing major clusters

like ‘RNA splicing’ and ‘mRNA processing’. This suggests that MDF has an important role in regulating the RNA splicing mechanism and this process might be carried out by means of RNA splicing itself. The second largest supercluster is ‘regulation of response to stress’. This is in line with the suggestion that alternative splicing is a crucial mechanism to response to developmental processes and environmental stress in literatures (Ling et al., 2017).

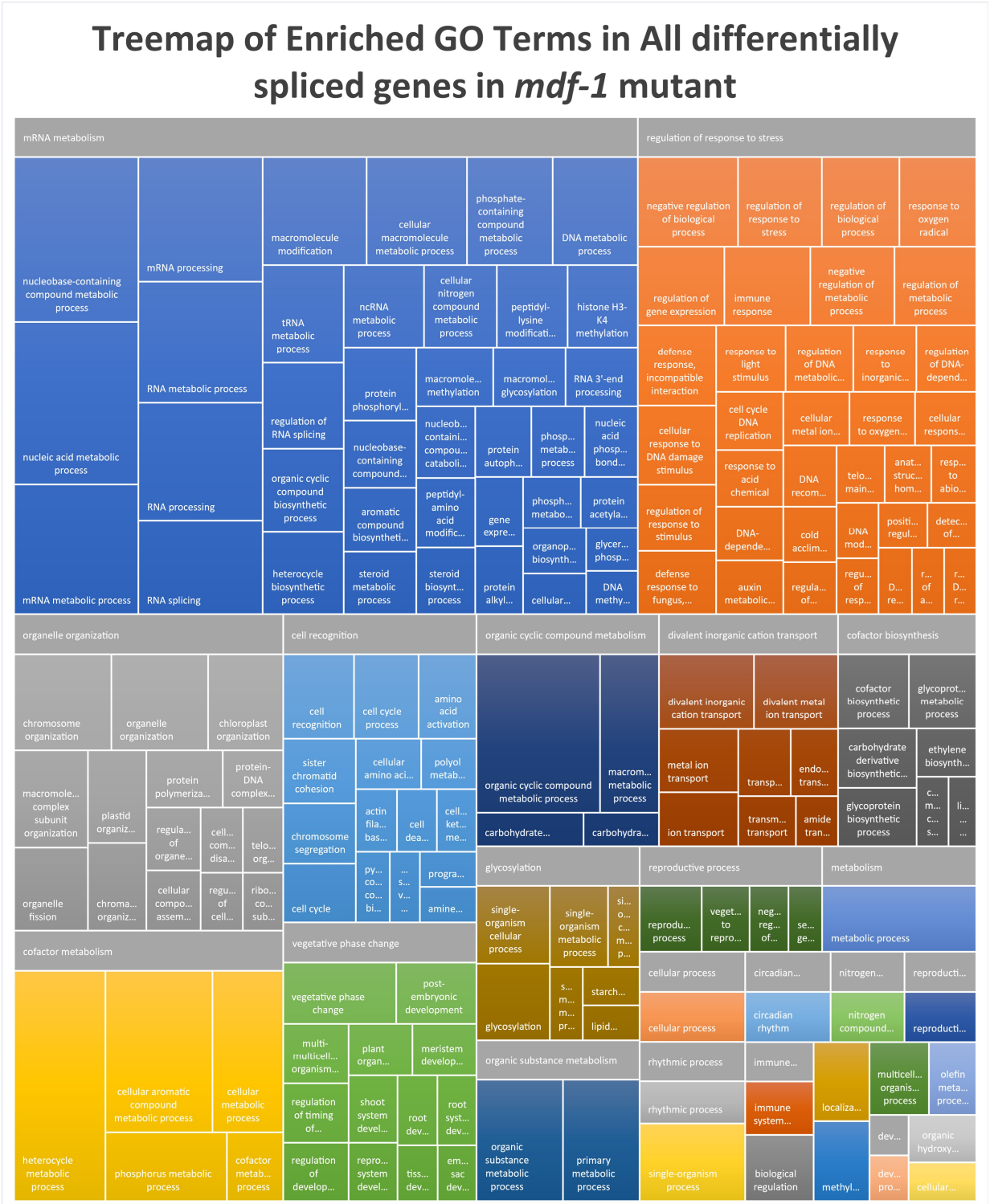


Figure 4.9. Tree map output from REVIGO showing the clustering of enriched gene ontology (GO) terms in the list of genes that are differentially spliced with FDR (false detection rate) < 0.01; ILD (Inclusion Level Difference) > 10% and < -10%. Each coloured rectangle represents a GO term cluster, and each colour with its grey bar on top and colour key below the tree map shows a supercluster of related clusters. The colour key is ordered by total size of the cluster from left to right on each row, then top down across. The size of each coloured rectangle represents the absolute value of the $\log_{10}$ P-value ($|\log_{10}$ q-value$|$) of its cluster.

4.3 Bioinformatic study on role of MDF on response to stress

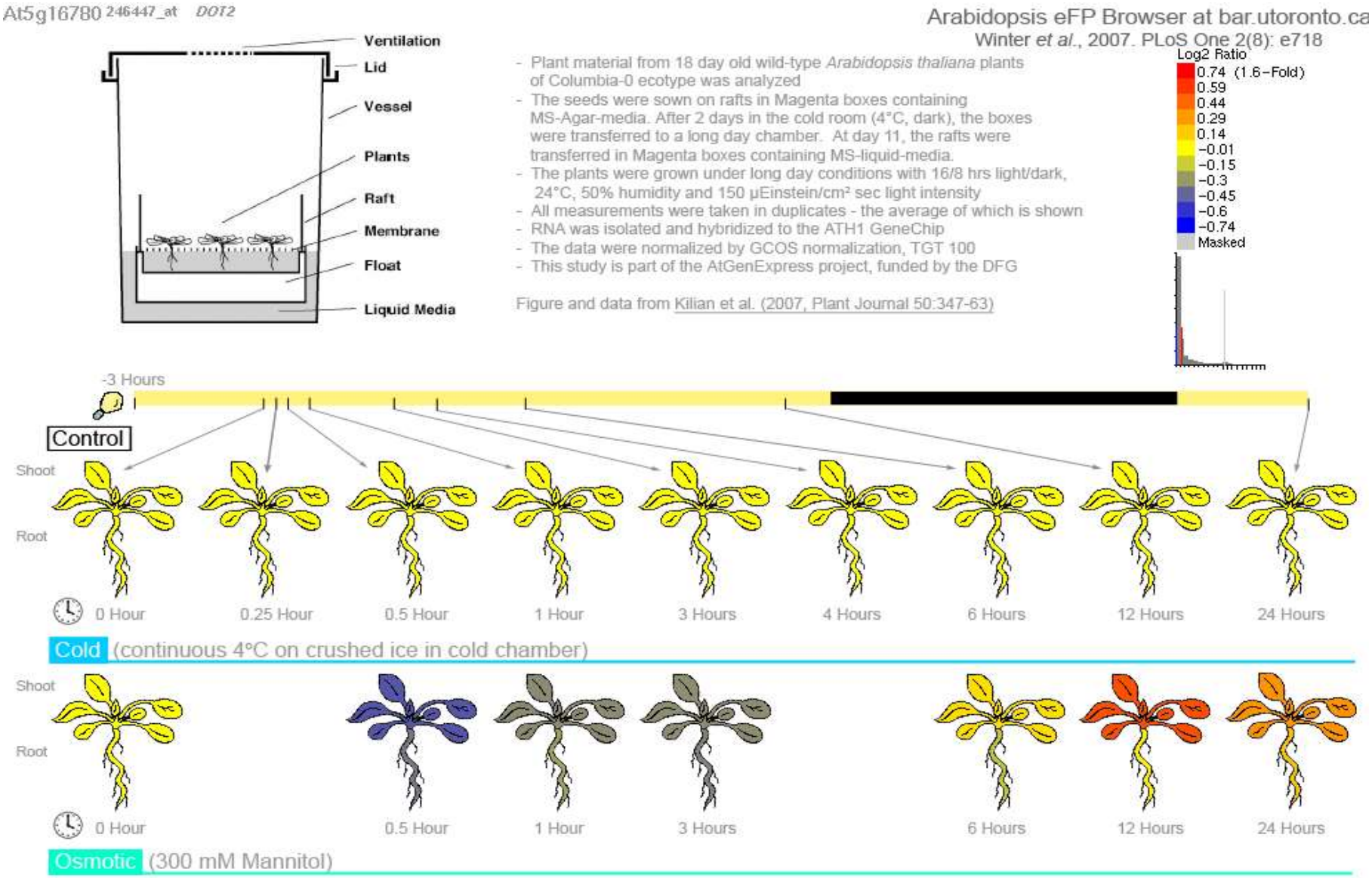


Figure 4.10 (1/3) Toronto eFP viewer on effects of various abiotic stresses (cold and osmotic stress) on expression level of *MDF* gene in *Arabidopsis*. Colour scale shows relative express level compared with control at T0.

Figure 4.10 (2/3) Toronto eFP viewer on effects of various abiotic stresses (salt, drought, genotoxic, and oxidative) on expression level of *MDF* gene in *Arabidopsis*. Colour scale shows relative express level compared with control at T0.

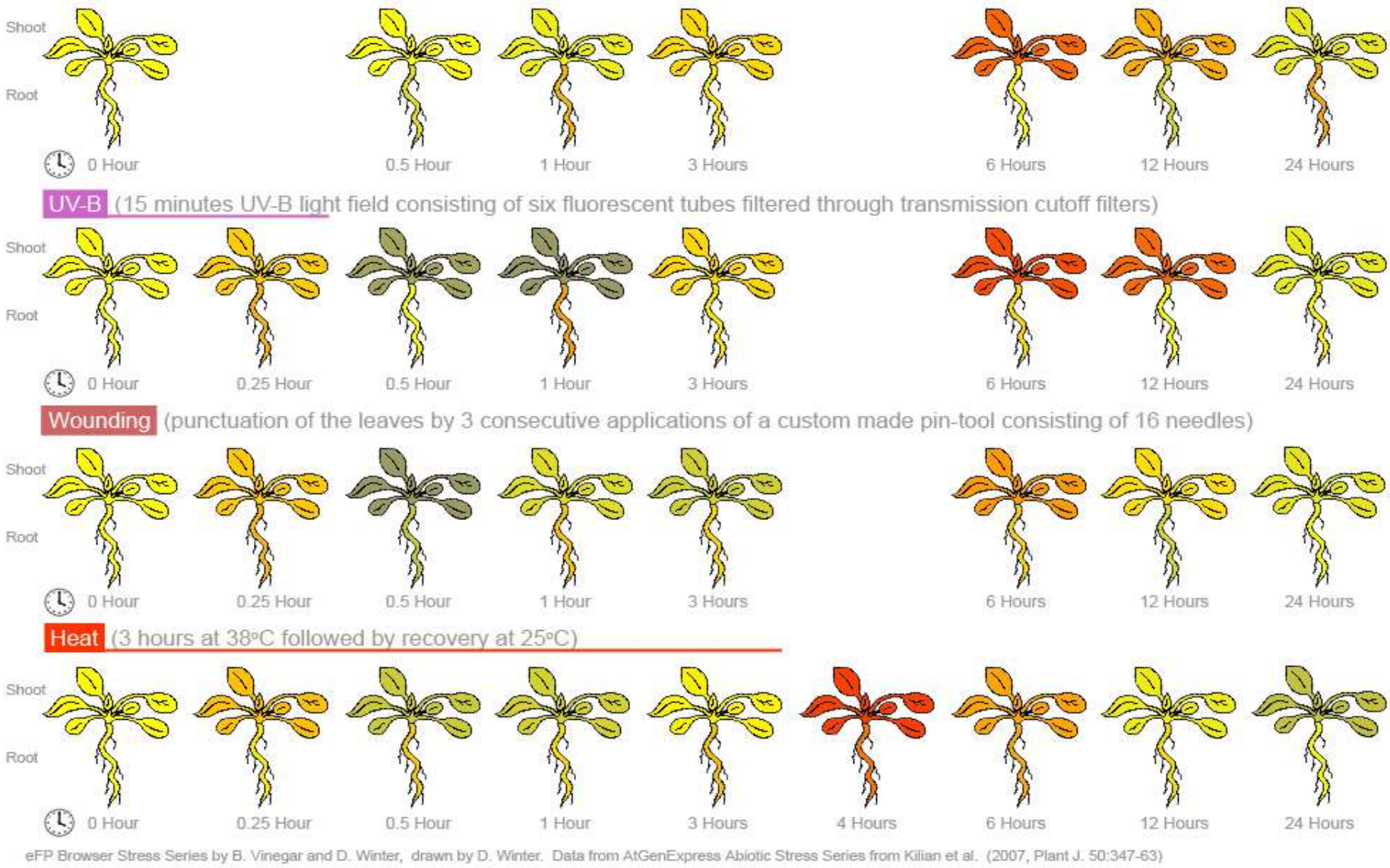


Figure 4.10 (3/3) Toronto eFP viewer on effects of various abiotic stresses (UV-B, wounding, and heat) on expression level of *MDF* gene in *Arabidopsis*. Colour scale shows relative express level compared with control at T0.

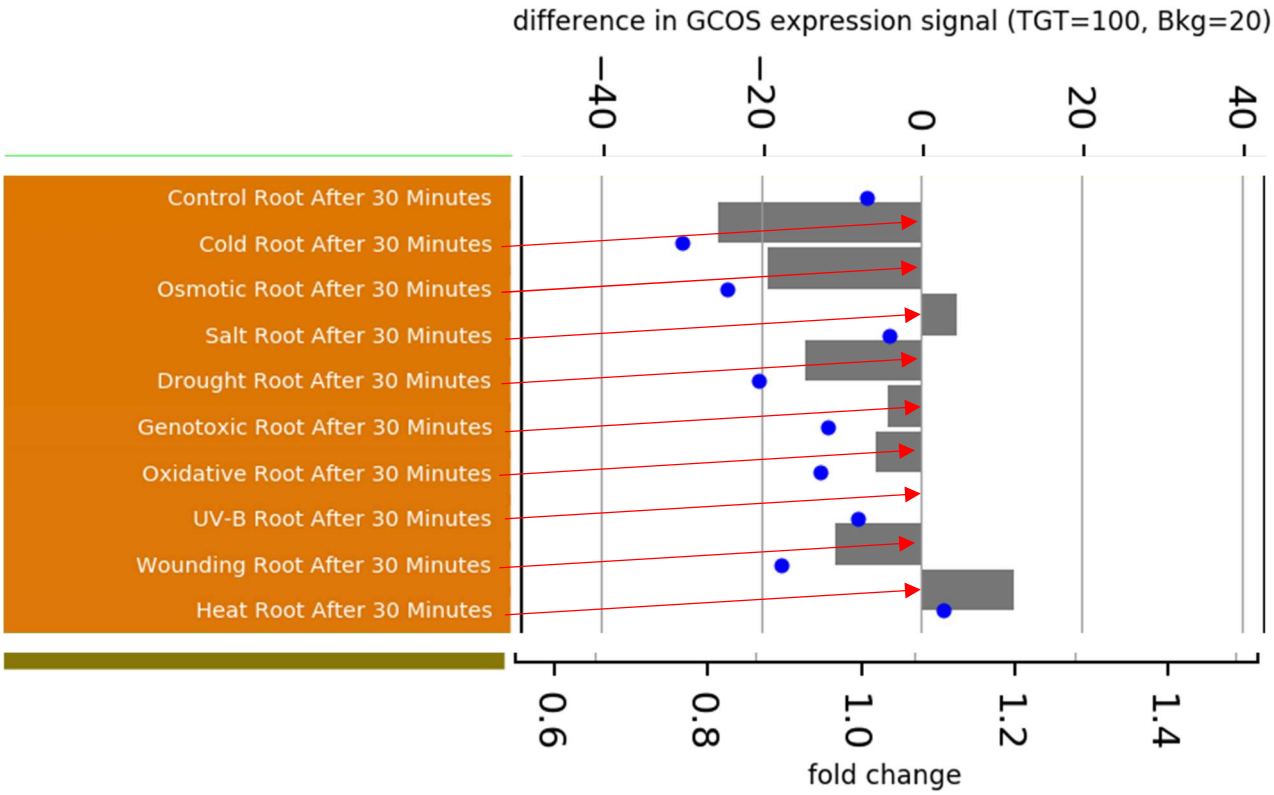


Figure 4.11. Bar chart showing relative expression level of *MDF* gene in *Arabidopsis* under different stress treatment after 30 minutes. Blue dots represent difference in GCOS expression signal compared with control measurement at T0; Grey bars show the fold change of each treated group reletave to the Control Root After 30 Minutes on the top.

Utilizing information from NASCArrays, the BAR eFP viewer from Toronto University provides a powerful to learn available data on gene expression across a variety of factors. A microarray study shows that *MDF* expression is significantly down regulated under a range of abiotic stresses. Along with the result from RNA sequencing data showing that significantly upregulated stress pathways in *mdf-1* transgenic lines, this negative correlation suggests that MDF plays an important role in regulating responses to environment stresses, likely through alternative splicing.

Summary

The study in this chapter helped improve understanding of how the transcriptome changed in response to mutation of the *MDF* gene in Arabidopsis, using data from Next Gen Sequencing technology. Gene ontology analysis using AgriGO and REVIGO on the RNA-seq data provided an overview on the changes, leading the way for more specific study using other bioinformatic and, in the future, experimental tools. The transcriptome analysis helped to confirm the hypothesis that MDF plays an important role in maintaining a normal meristematic development, very likely through means of regulating alternative splicing. The negative correlation between *MDF* expression and stress response suggests that MDF and its regulation in alternative splicing might also play an important role in regulating environmental adaptation and stress response.

Chapter 5 The IAA Biosynthetic Network

5.1 Introduction

The experimental results from previous chapters demonstrated that the POLARIS (PLS) peptide plays a crucial role in maintaining the homeostasis of ethylene and auxin signalling in Arabidopsis. More specifically, it shows that PLS is required for ethylene-mediated auxin biosynthesis in the root tip. Despite being the first plant hormone to be discovered, there is still not a clear picture of the biochemistry behind auxin biosynthesis (Mano and Nemoto, 2012). Many pathways have been postulated over the past decades, but none has proved to be indispensable so far. In order to understand how auxin biosynthesis may be regulated by POLARIS, it is important to understand how existing auxin biosynthesis pathways may connect to each other, and to identify which pathways may be affected by altered PLs levels. To achieve this, in this chapter I built a network of existing postulated auxin biosynthesis pathways using information from the literature and combined it with experimental data from my project.

Our group have recently proposed that POLARIS peptide is a Cu+ binding molecule that facilitate the ethylene signal transduction function of the ETR1 ethylene receptor (Mudge 2016 and Mudge *et al*. in preparation). This is consistent with the observed phenotype linked to the ethylene response in *pls* mutant seedlings (Chilley et al, 2006). However, the lack of ACC-induced auxin accumulation in the *pls* mutant root tip further confirmed the hypothesis that PLS might be playing an essential role in the auxin biosynthetic pathway independent of ethylene signalling.

Our group started modelling work by creating a hormonal crosstalk network, using experimental data and literature, to study how auxin concentration is controlled in a single Arabidopsis cell by the collective effects of auxin biosynthesis and transport, in conjunction with auxin, ethylene and cytokinin signalling and POLARIS peptide (Liu et al. 2010). It was then developed to include PIN1 and PIN2 activities (Liu et al. 2013) before moving on to a spatial model to show hormonal patterning across different developmental stages on a realistic root map showing individual cells (Moore et al. 2015, 2017) (Fig. 5.1). Recently the network was further improved by including abscisic acid signalling and effect of osmotic stress (Rowe et al., 2016). In this chapter, the auxin biosynthesis component of the network is expanded, using data from the literature and my own experiments, to help better understand the crosstalk between plant hormones and how they regulate root growth and development

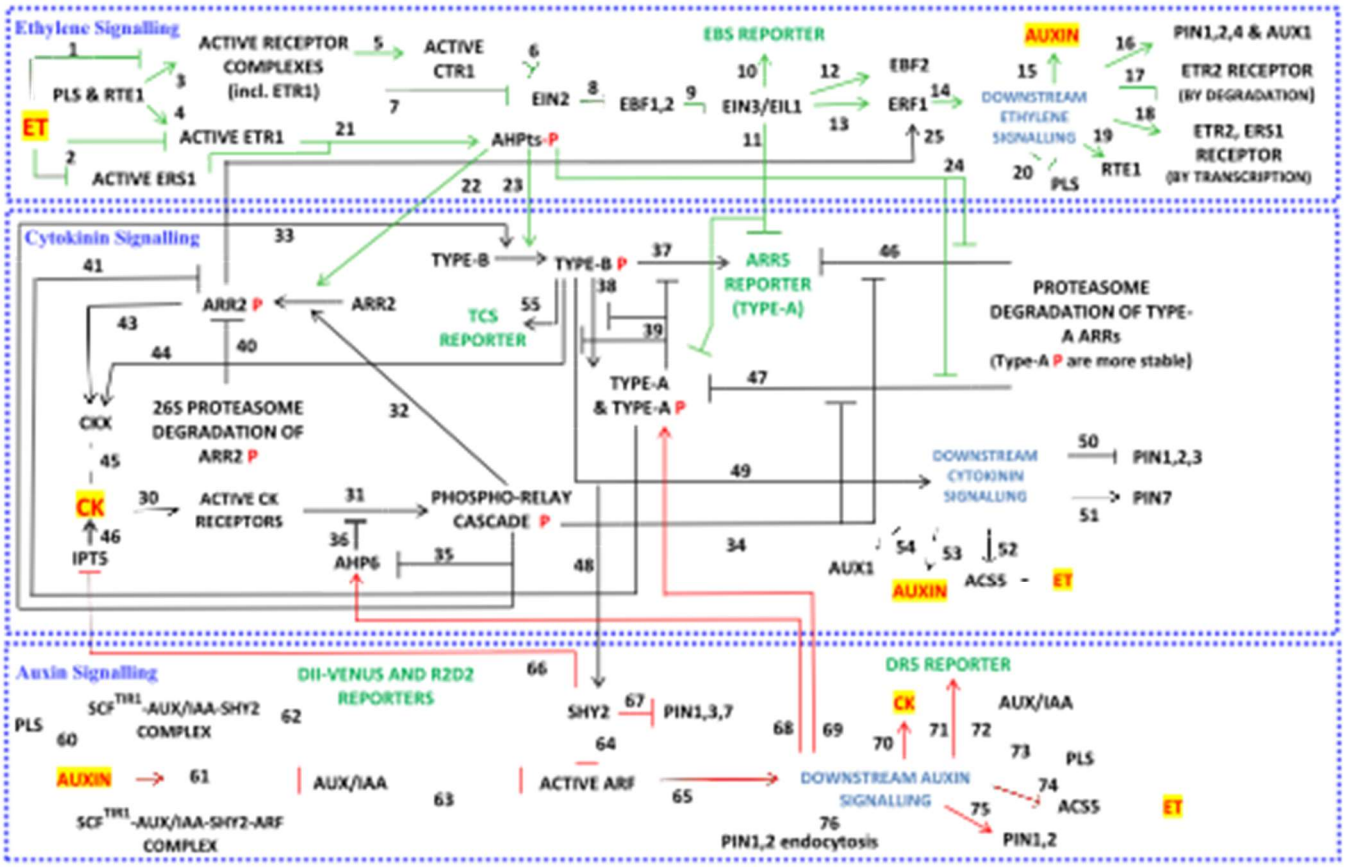


Figure 5.1 Integration of Experimental Data Reveals Multiple Layers of Complexity in Auxin, Cytokinin, and Ethylene Crosstalk in Arabidopsis Root Development.

Each number refers to a chemical reaction built into the model. Upper panel (green coloured links) schematically describes ethylene signalling pathways. Middle panel (black coloured links) schematically describes cytokinin signalling pathways. Lower panel (red coloured links) schematically describes auxin signalling pathways. A number by a link describes the link as summarized in Supplemental Table 1. The links connecting the three panels are the main crosstalk links between auxin, cytokinin, and ethylene. The three hormones are highlighted in yellow, and are placed in different locations in the three panels, further showing their crosstalk. / stands for positive regulation; –j stands for negative regulation.

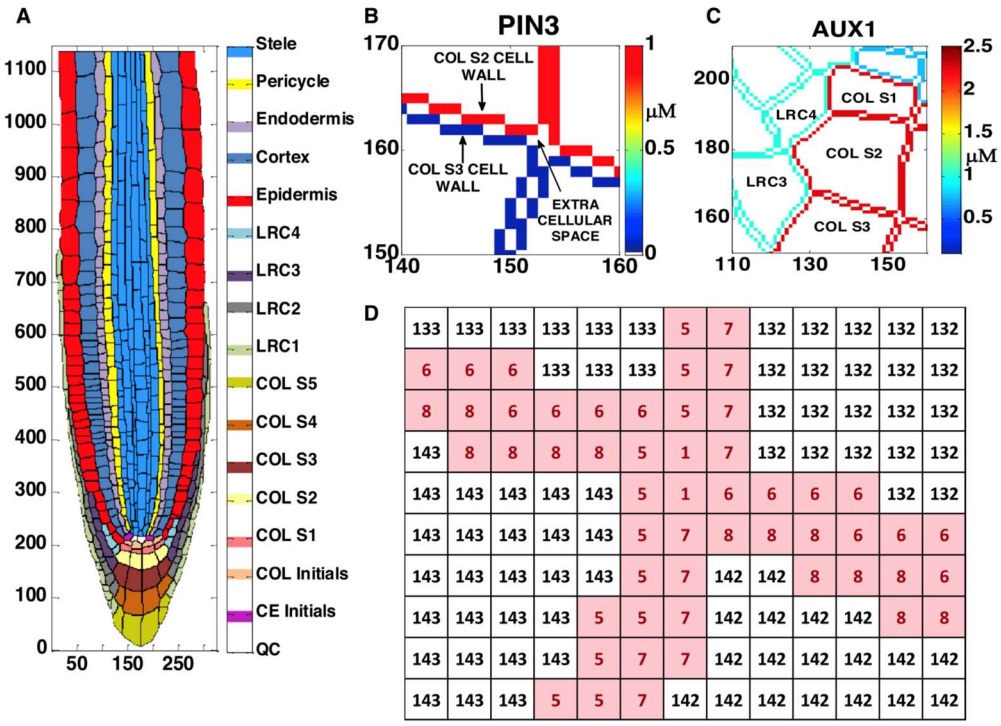


Figure 5.2 Construction of a Digital Root (Liu et al., 2017).
(A) A realistic root map showing the individual cells, based on confocal imaging. LRC 1 to 4, lateral root cap 1 to 4; COL S1 to S5, columella S1 to S5; CE initials, cortical endodermis initials; COL initials, columella initials; QC, quiescent centre.
(B) Localization of efflux (PIN3) carrier at the combined plasma membrane and cell wall entity of selected cells, with extracellular space between the cell walls of adjacent cells. COL S2 and S3, columella tier 2 and 3 cells.
(C) Localization of influx (AUX1) carrier at the combined plasma membrane and cell wall entity of selected cells, with extracellular space between the cell walls of adjacent cells. COL S1, S2, and S3, columella tier 1, 2, and 3 cells; LRC 3 and 4, lateral root cap tier 3 and 4 cells.
(D) A magnified part of the root to show an example of how to digitize the root. The root (A) can be discretized into grid points with any resolution (e.g., a grid point can be described by 2 mm multiplied by 2 mm in a two-dimensional space). A number is assigned to each grid point to describe the identity of this grid point. For the details of constructing a digital root, see Moore et al. (2015c, 2017). Numbers 132, 133, 142, and 143 are the grid points describing the cytosolic space of 132nd, 133rd, 142nd, or 143rd cell in the root, respectively. 1, 5, 6, 7, and 8 are used as ‘‘identifiers’’ to define grid points of the combined plasma membrane and cell wall entity or extracellular space, and they are also used to define distribution of both auxin efflux and influx carriers. Computational codes are used to calculate concentrations of all components in the hormonal crosstalk network at all grid points of the root.

5.2 Auxin Biosynthetic Pathways

Despite being the very first phytohormones being discovered, the exact mechanism of auxin biosynthesis in plant cells are still not clearly understood (Mano and Nemoto, 2012, Mashiguchi et al., 2011). Studies on auxin biosynthesis involve a broad range of plant species as well as bacteria, and there seems to be great redundancy among auxin biosynthesis pathways. The behaviours of many pathways are species specific, or even tissue specific within the same species (Kriechbaumer et al., 2016, Mashiguchi et al., 2011, Tivendale et al., 2014). Using available information from literature, I construct a network for IAA biosynthesis to include the most discussed postulations relevant to *Arabidopsis thaliana*.

The network is split into two parts depending on whether tryptophan (Trp) is involved as an intermediate. As the names suggests, the Trp-independent pathway synthesises IAA from indole, skipping tryptophan as an intermediate; while the Trp-dependent pathway incorporates tryptophan as part of the key ingredient for IAA biosynthesis.

Although Trp-dependent pathways have received most of the attention, even believed to be the only way that IAA can be synthesised (Tivendale et al., 2014). Experiments using isotopic tracer and metabolic analysis on *trp-1*, a loss-of-function mutant that cannot synthesise tryptophan, have confirmed the presence of at least some synthesised IAA without using exogenous tryptophan (Last and Fink, 1988). It is believed that Indole-3-glycerol phosphate or indole is the likely precursor. However, the exact mechanism of its biochemical pathway to IAA is still unknown.

There are mainly four Trp-dependent pathways proposed for *Arabidopsis thaliana* and they deviate at indole synthesis from the Trp-independent pathway. They are named after the intermediates immediate downstream of tryptophan - the IAM, IPA, IAOX and TAM pathways. There are varying degrees of evidence to support each proposed pathway, and as mentioned above, some of them are tissue-specific.

IAM Pathway

It has been confirmed that Trp is converted to IAM by iaaM (encoding the tryptophan-2-monooxygenase) in the bacteria *Agrobacterium tumefaciens* and *Pseudomonas savastanoi*, before being hydrolysed to IAA by iaaH (encoding IAM hydrolase) (Comai and Kosuge, 1982,

Yamada et al., 1985). This is proposed to be an important pathway for all plants, as IAM has been identified as an endogenous compound in many plant species including Arabidopsis, and is believed to be widespread across plant kingdom (Kriechbaumer et al., 2016). In Arabidopsis IAM is mainly synthesized from IAOx, which is part of a Brassica-specific auxin biosynthesis pathway, synthesized from Trp catalysed by the enzymes CYP79B2 and CYP79B3 (Sugawara et al., 2009). The lack of these enzymes and IAOx in other plant families fuelled speculation that IAM can also be generated by iaaM-like enzymes in plants. IAM then is converted to IAA by enzymes from the AMI1 family, which is believed to be widespread within the plant kingdom (Lehmann et al., 2010). To reflect these two distinctive sources of IAM, I added a separate IAM sub-pathway under IAOx pathway in my network.

Interest in the IPA pathway originated from studies on microorganisms that synthesize IAA (Koga et al., 1992, Koga, 1995). The IPA molecule is readily oxidised to IAA at room temperature (Koga et al., 1992). Due to its instability, it is particularly hard to characterize the enzymatic reactions of this pathway. As a result, there have only been suggestions that some YUCCA family proteins are playing a role in catalysing this reaction without knowing the exact mechanism (Dai et al., 2013). *TAA1*, encoding Trp aminotransferase responsible for converting Trp to IPA, was identified in a study of shade avoidance in *Arabidopsis*. The study showed that *taa1* seedlings have about 60% the level of IAA compared with WT IAA levels, and its IAA level does not change when being moved into shade as it does in WT. This suggests that the increased level of IAA exhibited in WT under shade is generated through this pathway, and it is likely that a sizeable proportion of IAA is synthesized through this pathway (Tao et al., 2008). There is also evidence for the activity of VAS1, which converts IPA and methionine, an ethylene precursor, to tryptophan and 2-oxo-4-methylthiobutyric acid, indicating that IPA plays a crucial role in coordinating the biosynthesis of ethylene and auxin (Zheng et al., 2013).

The IAOX Pathway

The IAOX pathway is believed to be a Brassicaceae-specific pathway for the biosynthesis of both IAA and glucosinolates (Ljung et al., 2005, Sugawara et al., 2009). It has been confirmed that the cytochrome p450 enzymes CYP79B2 and CYP79B3 catalyse the conversion of Trp to IAOX (Kriechbaumer et al., 2016). As with all the other postulated pathways, the exact biochemistry of conversion to IAA is still unclear. There is significant localisation of CYP79B2/3 in the root meristem region. Knocking out the function of both *CYP79B2/3* genes

had little effect on the overall concentration of IAA in Arabidopsis (Ljung et al., 2005, Sugawara et al., 2009). However, the double mutant had significantly reduced IAA biosynthesis levels in excised root tips, indicating that IAOX pathway might be responsible for IAA biosynthesis in the root meristem region (Ljung et al., 2005). In addition, *sur1* and *sur2* mutants both exhibit significantly elevated level of IAOX and IAA (Mikkelsen et al., 2004), further suggesting that a substantial amount of IAA can be converted from IAOX.

The TAM Pathway

The TAM pathway is possibly the best example of ambiguity in our understanding of the IAA biosynthesis pathways. It was originally proposed to play a role when TAM was observed to have auxin-like activities in an experiment with *Avena* coleoptiles in 1966 (Winter, 1966). Subsequently there were various papers showing that TAM is found in many plant species, including *Arabidopsis* (Cooney and Nonhebel, 1991, Sugawara et al., 2009). However, there were also results from radio-labelling metabolic experiments suggesting that the TAM and IAA synthesised in plants might not come from the same Trp pool, and TAM might not be a major source of IAA (Pollmann et al., 2002). Furthermore, accumulation of TAM in TDC- (Trp decarboxylase)-overexpressing tobacco plants did not correlate with significant changes in IAA levels (Songstad et al., 1990). Recent advances in the study of *YUCCA* gene families found that TAM might be one of the precursors of IAOX, either directly or through the synthesis of NHT, which was shown to be the *in vitro* product of TAM catalysed by YUCs (Zhao et al., 2001, Ljung, 2013). However, there is also evidence questioning the authenticity of NHT in the experiment (Tivendale et al., 2010). Thus far there is no definitive evidence to either confirm or reject the significance of the TAM pathway in IAA biosynthesis.

5.3 Construction of the IAA Biosynthesis Network

Using information from the literature alone, I constructed a network of IAA biosynthesis pathways (Fig. 5.4). Depending on the involvement of tryptophan, the network is separated into two boxes with dashed lines. Key intermediate compounds along all postulated pathways are shown in rectangles, colour coded for each pathway. Well recognised reactions are depicted using solid arrows with colour coding, accompanied by the

corresponding catalyst encoding gene(s). Arrows with dashed lines are used for proposed reactions that still trigger debates or lack evidence.

Incorporating PLS into the network

PLS is known to have negative regulatory effect on ethylene for a long time (Chilley et al 2006). Its mechanism was later confirmed to be a copper carrier for the ethylene receptor ETR1 (Mudge 2016 and Mudge et al., in preparation). With this evidence in mind, an inhibitory arrow was placed against the ethylene signalling pathway in the network (Fig. 5.4).

However, this arrow fails to explain why ACC-treated *pls* mutant seedlings do not exhibit increased local biosynthesis of auxin in the root tip like wild type seedlings do. Or, in other words, why is PLS required for ethylene-triggered auxin biosynthesis in the root tip? Upon looking closely at the RNA sequencing data for each postulated pathway, a pattern emerged.

Most of the genes shown in the network did not show significant change in expression level in the *pls* mutant RNA seq data (Chapter 3). However, 6 key genes in the IAOX pathway showed strong signals with high confidence level (Fig. 5.4). *CYP79B2/B3* and *SUR1/2* all showed significant downregulation in the *pls* mutant., while *CYP71A13* and *NIT2* both showed increased expression.

As mentioned above, the IAOX pathway is believed to be specific to the Brassicaceae, and there is evidence showing that it is responsible for IAA biosynthesis in the root tip (Kriechbaumer et al., 2016). GUS reporter promoter analysis of both *B2* and *B3* gene showed localisation pattern (fig 6.2) in the root tip meristem region , which convincingly overlaps with that of PLS.

Therefore, it is feasible that the lack of IAA biosynthesis in the *pls* mutant is caused by the downregulation of the IAOX pathway. The observed downregulation of B2/B3 gene in the *pls* mutant expression could lead to insufficient accumulation of IAOX , which is a common precursor of IAA and indole glucosinolates. To compensate for the lack of IAA biosynthesis, CYP71A13 and NIT2 are proposed to be both upregulated, maximising the use of limited IAOX. At the same time *SUR1/2* are down regulated to preserve IAOX from being used for synthesising indole glucosinolates.

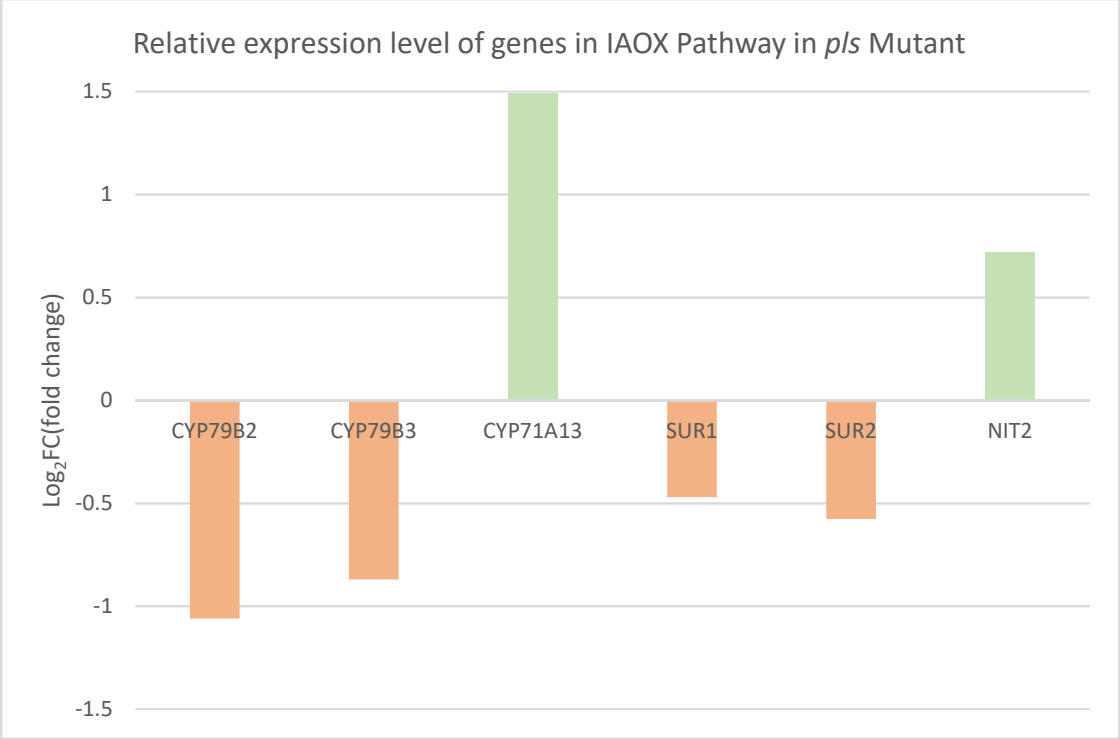


Figure 5.3. The relative expression level of key genes in the IAOX pathway plotted using RNA-seq data from *pls* mutant. Vertical axis represents log2FC (fold change), where +/- 1 represents two folds up/down regulation relative to that of wild type. q-value < 0.01 for all but SUR1, where q-value < 0.1.

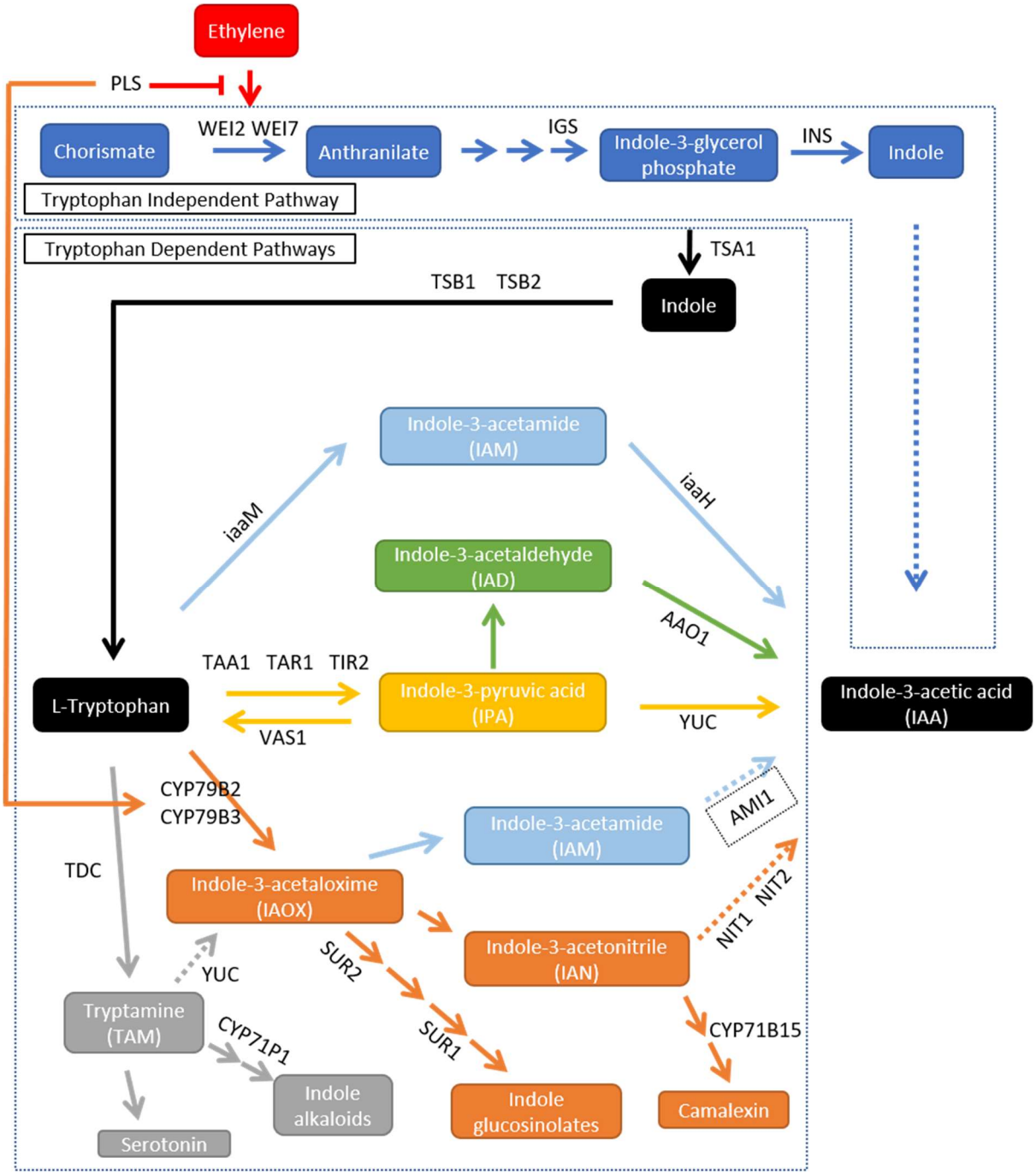


Figure 5.4 Auxin biosynthesis network constructed using postulated auxin biosynthesis pathways in the literature. The network is divided into two parts: tryptophan independent and dependent pathways. Trp dependent pathways are colour coded by intermediates. Each solid arrow represents a chemical reaction. Dotted arrows represent hypothetical reactions. Known genes are marked next to reaction arrows.

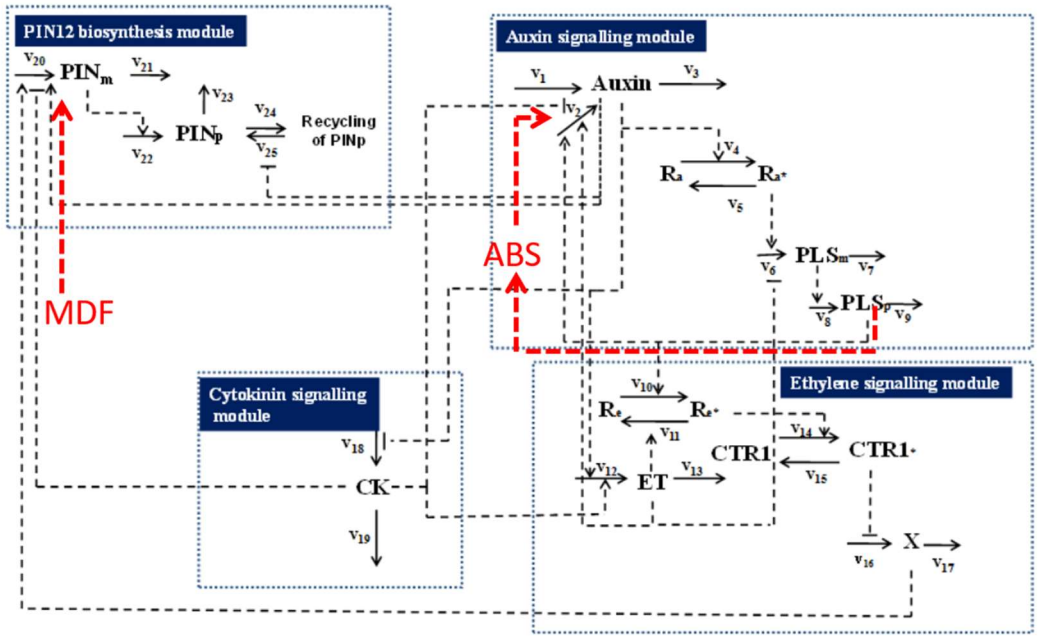


Figure 5.5 Revised hormonal crosstalk network. Using information from PLS and MDF RNA sequencing analysis, ABS (auxin biosynthesis switch) and MDF (in red) are added to the network from (Liu et al., 2017).

Summary

The aim of this chapter was to construct a new network of auxin biosynthesis pathways and incorporate into the existing hormone crosstalk model that includes a prospective role for PLS. The ambiguous nature of our understanding of auxin biosynthesis posed significant difficulties when constructing the network. Data from the literature were examined and tailored to construct the backbone of the network, and RNA seq data was used to add PLS into the network. Extreme care was taken when evaluating evidence from the literature, especially when there are opposing views, and this version of network accounts for these different possible pathway interactions.

The data from RNA sequencing experiments suggest that PLS is highly likely to be a positive regulator in the early part of IAOx pathway, potentially explaining the lack of IAA biosynthesis in the root tip of *pls* mutant seedlings. The network also shows the potential route by which ethylene may influence auxin biosynthesis, in PLS-dependent and -independent ways.

Chapter 6. Discussion

Auxin is one of the major phytohormones that affect many aspects of plant development and growth, and it plays a key role in the interactions between phytohormones and regulatory genes in plants. The importance of understanding phytohormone cross talk is far more important than merely fulfilling the great curiosity of humankind. It has practical implications in areas including environmental protection, agricultural optimisation, and bioenergy production.

In this thesis, RNA sequencing experiments were carried out on *pls* and *mdf* mutant seedlings to facilitate the study of these genes' potential roles in auxin homeostasis. Previous work had suggested these genes encode proteins essential for this process, but the mechanisms involved were unclear (Chilley et al. 2006; Casson et al. 2002, 2009). To separate auxin transport from local biosynthesis, NPA was used in bioimaging experiments for its auxin-transport blocking properties (Sabatini et al. 1999). Bioinformatics analysis was also used to further investigate the roles of PLS and MDF. Building gene-signalling networks using data from the literature and experiments described in this thesis helped to create a revised overview of our current knowledge of auxin biosynthesis pathways.

6.1 The POLARIS peptide is required for ethylene signalling control

The data from my RNA sequencing experiments show that many of the top 20 enriched gene ontology terms in the list of upregulated genes in *pls* mutant seedlings are associated with ethylene-related biological processes. This supports previous observations that the *pls* seedlings exhibit enhanced ethylene signalling, causing its indicative short-root phenotype (Chilley et al., 2006); but the data provide a high resolution analysis of the transcriptional changes associated with loss of PLS function. This finding is also in line with the hypothesis on the close relationship between PLS and ethylene, which our group has recently confirmed by the identification of POLARIS peptide's function in ethylene signalling, where it acts as a metallochaperone, donating a copper (I) ion to the ethylene receptor ETR1 (Mudge, 2016). This Cu ion is essential for correct ethylene binding of ethylene molecule and transduction of its signal (Rodriguez et al., 1999).

Working with other hormones, ethylene plays a crucial role in plant growth and development, including regulation of cell elongation, differentiation, cell death, tissue patterning, root development, gravitropism and response to stress (Van de Poel et al. 2015).

This finding further confirms the hypothesis that PLS modulates a wide range of cellular functions and activities through negatively regulating ethylene signalling.

6.2 The POLARIS peptide has a role in regulating photosynthesis genes

Among the thousands of genes that show altered expression levels in the *pls* mutant and the PLSox transgenic overexpresser, only 212 are up-regulated in *pls* as well as down-regulated in *PLSox*. Further gene ontology analysis on these genes showed that the most prominent biological processes these genes are involved with are all related to photosynthesis. There is no previous information in the literature on the relationship between the PLS peptide and photosynthesis. However, GUS staining in transgenics containing a pPLS::GUS promoter fusion showed that *PLS* expression occurs predominantly in non-photosynthetic tissues, including the root meristem region and leaf vascular tissue (Casson et al., 2002). This suggests that PLS might be acting as a negative regulator on the differentiation of photosynthetic tissue, possibly through the complex network of hormones controlling cell differentiation, or through its metal binding activity directly controlling the activity of relevant enzymes or other proteins involved in chlorophyll biosynthesis, for example.

6.3 The POLARIS peptide is essential for ethylene-induced auxin biosynthesis in the Arabidopsis root tip

The inhibitory effect on root growth of ethylene signalling is achieved through the accumulation of auxin in the root tip meristem region, and its transport to the elongation zone of the root due to enhanced PIN gene expression (Ruzicka et al. 2007; Swarup et al. 2007). The imaging experiment (Fig. 3.16) confirmed this effect using exogenous application of the ethylene precursor ACC, which induced elevated auxin accumulation in the root tip of wild type Arabidopsis seedlings. The lack of this effect in the *pls* mutant strongly suggests that the PLS peptide is required for ethylene-induced auxin accumulation in the root tip.

The two main sources of auxin in root tip are local biosynthesis and auxin transport down the shoot (Chilley et al., 2006). The bioimaging analysis on NPA-treated seedlings (Fig. 3.16) ruled out the effect of auxin polar transport (which potentially could have removed auxin from the root tip to generate the observed low auxin content of the *pls* root tip (Chilley et al. 2006). This further supports the hypothesis that, in addition to facilitating the generation of a functional ethylene receptor ETR1, PLS is also directly involved in regulating auxin

biosynthesis, independent of ethylene signalling, at least in the root meristem region. The mechanism of such regulatory activity is yet to be determined. However, considering that there is now strong evidence that PLS facilitates copper (I) transfer to ETR1, it seems feasible that a similar metallochaperone mechanism is responsible for regulating auxin biosynthesis, by providing Cu or some other metal cofactor to key enzymes in auxin biosynthesis; i.e., by a post-translational mechanism. This mechanism is proposed to work like a switch, whereby the PLS peptide is required for the switch to be activated. For the ease of discussion, this thesis refers it as an 'auxin biosynthesis switch' (ABS), and is discussed further below.

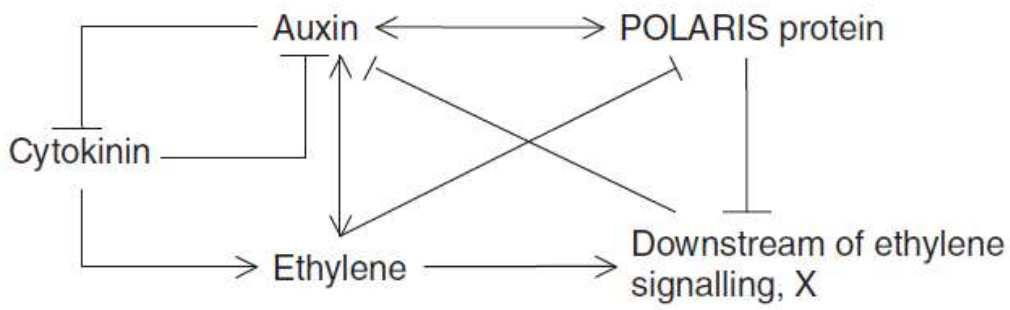


Figure 6.1 (Liu et al., 2010) The early hormonal crosstalk network revealed by modelling and experimental analysis in *Arabidopsis*.

This finding, of the putative role for PLS in auxin biosynthesis control, also sheds some light on the continuous development of our hormone crosstalk network (Fig. 6.1), further expanding our knowledge of how hormones and key genes interact with each other to regulate plant growth and development. The ABS that is mediated by PLS to control ethylene-induced auxin biosynthesis could in principle be logically placed at the position of molecule(s) X in this network, which was introduced in the original model to represent a group of unknown molecules facilitating the interaction between key components of the network, namely ethylene, auxin, and POLARIS (Figure 6.1). In other words, ethylene signalling could involve PLS to control ethylene-mediated auxin biosynthesis, as the work in this thesis demonstrates this essential role of PLS. However, ABS differs from molecule X in its relationship with PLS peptide and auxin. In the original model in Fig. 6.1, molecule X is inhibited by PLS and it is inhibitory to auxin, whereas ABS is positively regulated by POLARIS, and itself is a positive regulator of auxin biosynthesis. In this sense, it is desirable to put ABS

as a new component in the network, and a revised network needs to include this additional role for PLS.

6.4 The IAOX pathway is likely to be the main POLARIS-regulated auxin biosynthesis pathway in Arabidopsis root tip

The network of postulated IAA biosynthesis pathways, described in the previous chapter, gives a comprehensive overview of the current consensus in the literature on our understanding of the biochemistry underpinning IAA biosynthesis. By combining our RNA sequencing data with this analysis, we can propose how the lack of PLS peptide impacts the network, and this study suggests that the IAOX pathway is the most likely candidate for the ethylene-induced auxin biosynthesis that requires the presence of PLS.

In this Brassicaceae-specific pathway, IAOX is produced from Trp by CYP79B2/3 proteins. It has been demonstrated that the loss of function *cyp79b2/3* mutant failed to produce IAOX in Arabidopsis (Sugawara et al., 2009). Expression studies using promoter::GUS constructs showed that both *CYP79B2/3* genes have localisation patterns in the root meristem region that overlap with *PLS* expression (Figure 6.2) (Ljung et al., 2005). In the RNA sequencing data from the *pls* mutant, *CYP79B2/3* genes showed significant downregulation compared with wild type, suggesting that it is a rate limiting step in the IAOX pathway downstream of *PLS* function. In other words, PLS function is necessary for the expression of these genes and, by implication, the functioning of the IAOX pathway for auxin biosynthesis.

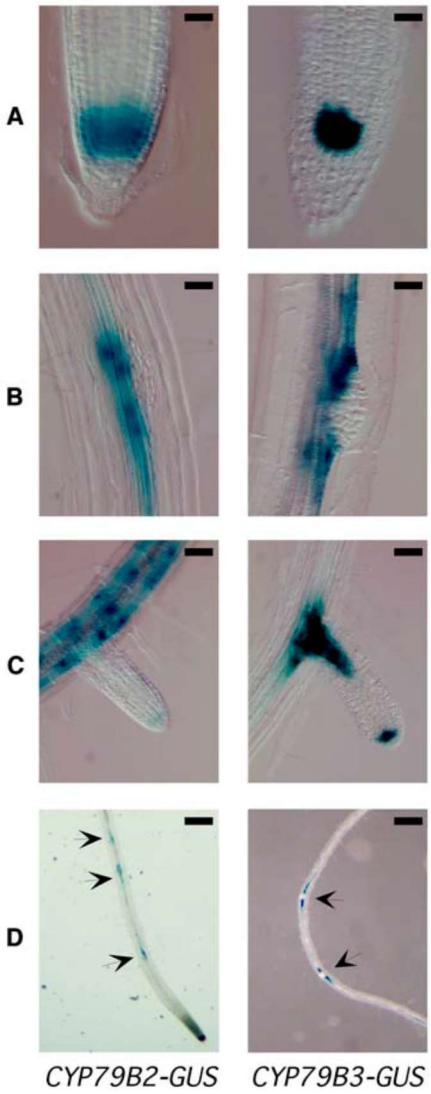


Figure 6.2 CYP79B2-GUS and CYP79B3-GUS Localisation in Arabidopsis Roots (Ljung et al., 2005).
CYP79B2-GUS transgenic plants are shown at left, and CYP79B3-GUS transgenic plants are shown at right. CYP79B2-GUS plants were stained with X-Gluc for 1 h, and CYP79B3-GUS plants were stained for 18 h.
(A) to (C): Expression patterns of CYP79B2 and CYP79B3 in the primary root tip (A), sites of lateral root formation (B), and developing lateral roots (C) of 7-DAG seedlings.
(D): The primary root after treatment with 1 mM IAA for 24 h at 7 DAG, showing expression at all sites of lateral root formation (arrows) but not in the rest of the primary root. Bars ¼ 25 mm in (A) and (B), 50 mm in (C), and 250 mm in (D).

Another important function of IAOX in the Brassicaceae is to produce indole glucosinolate through genes including *SUR1* and *SUR2*, and this compound appears to share the common pool of IAOX with auxin biosynthesis. Knocking out the function of the *SUR1* or *SUR2* genes results in significantly increased level of IAOX and auxin accumulation in seedlings, having a similar phenotype to that of wild type seedlings treated with exogenous auxin, with ectopic roots (Boerjan et al., 1995). Interestingly, genetic crosses have shown that the *pls* mutation acts as a phenotypic suppressor of the *superroot* mutant, i.e. introduction of the *pls* mutation reduces the frequency of ectopic roots in the *sur* mutant, presumably through a reduction of auxin content characteristic of *pls* (Casson et al., 2002). The RNA seq data show that both *SUR1* and *SUR2* are significantly downregulated in the *pls* mutant. A plausible interpretation would be that the system is trying to compensate the lack of IAOX for auxin biosynthesis, by a feedback mechanism. It can be speculated that evolution has deemed that auxin biosynthesis has higher priorities over synthesis of glucosinolates, which are secondary

metabolites associated with defence. It would be interesting to look further and possibly measure the effect of the *pls* mutation on glucosinolate content. In response to reduced auxin levels, the upregulation of the *CYP71A13* and *NIT2* genes follow a similar pattern, to boost auxin biosynthesis pathway.

6.5 MDF is required for correct RNA splicing

Alternative splicing enables organisms to generate more transcripts from a much smaller number of genes to increase diversity in the transcriptome and proteome. This ancient mechanism is highly conserved, and is shared by all eukaryotes (Black, 2003). The MDF human homologue, the hSART-1 protein, is known to be essential for the assembly of the spliceosome, and its absence is inhibitory to pre-RNA splicing activity (Gottschalk et al., 1999, Makarova et al., 2001). A prominent pattern in the RNA sequencing data obtained for the *mdf-1* mutant is the thousands of differential splicing events, where transcripts are not spliced properly compared to the wildtype. This is consistent with the possibility raised by Casson et al. (2009) that MDF may serve as a key component in the pre-mRNA splicing mechanism. The defect in splicing of so many genes seen in two independent *mdf* mutants is most likely the cause of the dramatic change in the transcriptome, leading to the severe phenotypic abnormalities seen in the *mdf* mutant meristems.

6.6 MDF is required for correct auxin patterning and meristem development

One of the key features of *mdf* mutant seedlings is the failure to maintain the quiescent centre (QC) and surrounding stem cells, leading to disrupted meristem patterning. Meristem organisation and development requires a stable auxin gradient and concentration maximum around the QC, which are delicately maintained by auxin transport proteins and local auxin biosynthesis (Sabatini et al., 1999; Aida et al. 2004). Members of the PIN protein family are responsible for the polar transport of auxin across membrane to the adjacent cell (Adamowski and Friml 2015).

Data from the RNA sequencing experiment on *mdf* mutants revealed that the transcript levels of *PINs1-7* in both *mdf-1* and *mdf-2* are significantly downregulated with high statistical confidence level. This shows that MDF plays a key role in regulating the expression the PIN protein family. Since there is no evidence that these *PIN* transcripts are mis-spliced in the *mdf* mutants, it is likely that in these mutants, one or more key genes, like

transcription factors that are responsible for regulating PIN protein family, are mis-spliced, causing the reduced PIN gene expression and leading to the observed disruption in polar auxin transport (Casson et al. 2009). Without the necessary auxin gradient and maximum at the root tip, the key meristem forming genes fail to establish the QC to form a functional meristem. Considering that the *plt1* and *plt2* double mutant does not show significantly disrupted auxin patterning, it is likely that the *PLT* genes are down stream of auxin signalling (Aida et al., 2004). As a result, it is hard to elucidate whether the reduced expression of the *PLT* genes in the *mdf* mutants is the product of differential splicing independent of disrupted auxin accumulation, or due to the reduced auxin maximum. It is likely that there is regulatory loop involving MDF, PLTs and PINs to control QC identity and meristem pattern and activity.

6.7 MDF has a role in regulating responses to environment stresses.

The RNA sequencing data show that many key stress-related genes are upregulated in *mdf* mutants. In addition, gene ontology analysis suggests that stress-related terms are among the most prominent. While this is evidence that MDF is a suppressor of stress responses and cell differentiation, further study was carried out to investigate this possibility. This led to a bioinformatics study on *MDF* expression levels in response to environmental stresses (Chapter 4). The aggregated microarray data from NASCArrays shows that MDF transcription exhibits a rapid response to many environment stimuli. 30 minutes after each of a range of environmental stress treatments, including cold, osmotic stress, drought, genotoxin, and wounding, the expression level of MDF showed significant downregulation in the root, while salt and heat treatment results in slightly increased MDF expression level. This shows that MDF is among the fast reacting genes that are quickly adjusted when plant need to respond to sudden change in surroundings, suggesting that MDF may play a key role in regulating stress related pathways. Given it is also required for maintaining stem cell identity in the root meristem, a model can be proposed in which MDF plays a key role in regulating the balance between two antagonistic processes, namely the maintenance of stem cell behaviour versus the activation of stress responses, secondary metabolism and cell differentiation and (ultimately) cell death (Figure 6.2). The reduction in *MDF* expression in response to environmental stress potentially provides a mechanism to restrict growth through inhibition of cell division as root auxin transport and levels are reduced in the meristem.

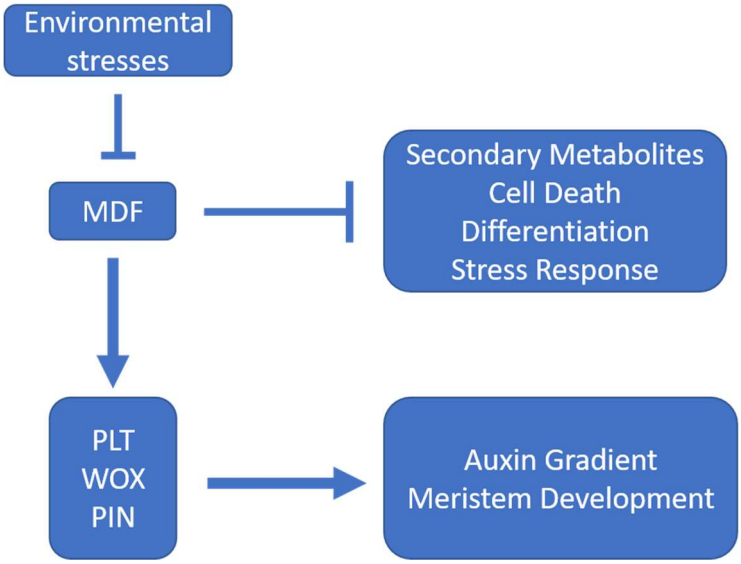


Figure 6.3 MDF responds to environmental stresses and regulates a variety of genes and biological functions.

6.8 Future work

6.8.1 ABS molecule in kinetic model

This thesis found that in the Arabidopsis root, the IAOX pathway is responsible for auxin biosynthesis, which is controlled by the PLS peptide. However, it is still unclear how PLS controls the expression levels of *CYP79B2/3* genes, or the activity of the encoded proteins. Further work is required to identify the ABS components to better understand the controlling mechanism. Potentially by more in depth analysis using existing rich collection of RNA sequencing data, candidate genes or pathways related to *CYP79B2/3* regulation can be identified, which in turn can lead to more exploration and confirmatory work Combining currently knowledge in the mechanism of PLS in ethylene signally, discovering its activities in other pathways could shed more light on how small peptide carry out regulatory activities.

6.8.2 Confirming the IAOX pathway by measuring concentration of IAOX and glucosinolate in *pls* mutant

Apart from auxin biosynthesis, IAOX is also required for glucosinolate synthesis. Available data show that when IAOX is in short supply, glucosinolate synthesis is tuned down, giving priority to auxin biosynthesis.

To confirm the hypothesis that the lack of auxin in ACC-treated *pls* seedling root tip is also due to reduced IAOX activity, it would be helpful to compare the IAOX and glucosinolate contents using metabolic profiling. If the hypothesis is true, *pls* mutant seedlings treated with ACC are predicted to have lower concentrations of IAOX and glucosinolate compared with ACC-treated wild type seedlings. If confirmed, this would further solidify the importance of IAOX pathway in Arabidopsis, auxin biosynthesis pathways, and more interestingly, further expand the hormonal crosstalk model to include secondary metabolism in plants.

6.8.3 PLS and Photosynthesis

Analysis of RNA sequencing data showed direct correlation between gene expression levels between *PLS* and photosynthesis-related genes. However, there are little data or literature linking these two. More work can be done to gain more knowledge in development of photosynthetic tissues and how it is regulated by PLS, such as measuring photosynthetic activities of *pls* and *PLSox* transgenic lines more directly, such as measuring uptake of carbon dioxide, production of oxygen, or production of carbohydrates (increase in dry mass). New findings will expand our knowledge in small peptide signalling and regulation of tissue specific development and growth, and potentially adding regulation of photosynthesis activities into the hormonal crosstalk model.

6.8.4 Validation of abnormal splicing events discovered in RNA sequencing experiment.

The data from RNA sequencing experiment gave an overview of alternative splicing events in *mdf-1/2*. However, when it comes to investigating individual genes, quantitative PCR can normally give much better resolution. Therefore, it would be useful to identify key genes that are shown wrongly spliced in *mdf* mutants, and validate using qPCR. A better understanding of alternative splicing control mechanism opens more opportunity in leaning plant growth and development, and their response to external changes in the environment.

6.9 Summary

In summary, the work described in this thesis provides new information on the mechanisms of hormone signalling in relation to root growth and development. The studies on two genes expressed in the root tip, *PLS* and *MDF*, provide new insight into the complexities of gene-hormone signalling networks linked to the control of cell identity, cell division, cell elongation and cell differentiation, which must all be balanced to construct a functional root meristem. The role of auxin is central to this study, but it is important to recognise that it does not function in isolation, but as one component of a complex network of signalling pathways that exhibit crosstalk. Future studies require a combination of experimental and predictive modelling approaches to further dissect control mechanisms.

Bibliography

AIDA, M., BEIS, D., HEIDSTRA, R., WILLEMSEN, V., BLILOU, I., GALINHA, C., NUSSAUME, L., NOH, Y. S., AMASINO, R. & SCHERES, B. 2004. The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche. *Cell,* 119**,** 109-20.

ALBERTS B, J. A., LEWIS J, ET AL. 2002. Signaling in Plants. *Molecular Biology of the Cell.* 4th edition ed.: New York: Garland Science.

ANDREW, S. 2016. FastQC - A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

ARABIDOPSIS GENOME, I. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature,* 408**,** 796-815.

BLACK, D. L. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem,* 72**,** 291-336.

BLEECKER, A. B. & KENDE, H. 2000. Ethylene: a gaseous signal molecule in plants. *Annu Rev Cell Dev Biol,* 16**,** 1-18.

BOERJAN, W., CERVERA, M. T., DELARUE, M., BEECKMAN, T., DEWITTE, W., BELLINI, C., CABOCHE, M., VAN ONCKELEN, H., VAN MONTAGU, M. & INZE, D. 1995. Superroot, a recessive mutation in Arabidopsis, confers auxin overproduction. *Plant Cell,* 7**,** 1405-19.

BOHLMANN, J., LINS, T., MARTIN, W. & EILERT, U. 1996. Anthranilate synthase from Ruta graveolens - Duplicated AS alpha genes encode tryptophan-sensitive and tryptophan-insensitive isoenzymes specific to amino acid and alkaloid biosynthesis. *Plant Physiology,* 111**,** 507-514.

CASSON, S. A., CHILLEY, P. M., TOPPING, J. F., EVANS, I. M., SOUTER, M. A. & LINDSEY, K. 2002. The POLARIS gene of Arabidopsis encodes a predicted peptide required for correct root growth and leaf vascular patterning. *Plant Cell,* 14**,** 1705-1721.

CASSON, S. A., TOPPING, J. F. & LINDSEY, K. 2009. MERISTEM-DEFECTIVE, an RS domain protein, is required for the correct meristem patterning and function in Arabidopsis. *Plant J,* 57**,** 857-69.

CHANDLER, J. W. 2009. Local auxin production: a small contribution to a big field. *Bioessays,* 31**,** 60-70.

CHILLEY, P. M., CASSON, S. A., TARKOWSKI, P., HAWKINS, N., WANG, K. L. C., HUSSEY, P. J., BEALE, M., ECKER, J. R., SANDBERG, G. K. & LINDSEY, K. 2006. The POLARIS peptide of Arabidopsis regulates auxin transport and root growth via effects on ethylene signaling. *Plant Cell,* 18**,** 3058-3072.

COMAI, L. & KOSUGE, T. 1982. Cloning characterization of iaaM, a virulence determinant of Pseudomonas savastanoi. *J Bacteriol,* 149**,** 40-6.

COONEY, T. P. & NONHEBEL, H. M. 1991. Biosynthesis of indole-3-acetic acid in tomato shoots: Measurement, mass-spectral identification and incorporation of (-2)H from (-2)H 2O into indole-3-acetic acid, D- and L-tryptophan, indole-3-pyruvate and tryptamine. *Planta,* 184**,** 368-76.

DAI, X., MASHIGUCHI, K., CHEN, Q., KASAHARA, H., KAMIYA, Y., OJHA, S., DUBOIS, J., BALLOU, D. & ZHAO, Y. 2013. The biochemical mechanism of auxin biosynthesis by an arabidopsis YUCCA flavin-containing monooxygenase. *J Biol Chem,* 288**,** 1448-57.

DAVIES, P. J. 2010. *The Plant Hormones: Their Nature, Occurrence, and Functions,* Dordrecht, Springer.

DE SMET, S., CUYPERS, A., VANGRONSVELD, J. & REMANS, T. 2015. Gene Networks Involved in Hormonal Control of Root Development in Arabidopsis thaliana: A Framework for Studying Its Disturbance by Metal Stress. *Int J Mol Sci,* 16**,** 19195-224.

FARROKHI, N., WHITELEGGE, J. P. & BRUSSLAN, J. A. 2008. Plant peptides and peptidomics. *Plant Biotechnol J,* 6**,** 105-34.

FORESTAN, C. & VAROTTO, S. 2012. The Role of PIN Auxin Efflux Carriers in Polar Auxin Transport and Accumulation and Their Effect on Shaping Maize Development. *Molecular Plant,* 5**,** 787-798.

GALINHA, C., HOFHUIS, H., LUIJTEN, M., WILLEMSEN, V., BLILOU, I., HEIDSTRA, R. & SCHERES, B. 2007. PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature,* 449**,** 1053-7.

GOTTSCHALK, A., NEUBAUER, G., BANROQUES, J., MANN, M., LUHRMANN, R. & FABRIZIO, P. 1999. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *EMBO J,* 18**,** 4535-48.

HAECKER, A., GROSS-HARDT, R., GEIGES, B., SARKAR, A., BREUNINGER, H., HERRMANN, M. & LAUX, T. 2004. Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana. *Development,* 131**,** 657-68.

HELARIUTTA, Y., FUKAKI, H., WYSOCKA-DILLER, J., NAKAJIMA, K., JUNG, J., SENA, G., HAUSER, M. T. & BENFEY, P. N. 2000. The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling. *Cell,* 101**,** 555-67.

JIAN, O. Y., SHAO, X. & LI, J. Y. 2000. Indole-3-glycerol phosphate, a branchpoint of indole-3-acetic acid biosynthesis from the tryptophan biosynthetic pathway in Arabidopsis thaliana. *Plant Journal,* 24**,** 327-333.

KATEKAR, G. F. & GEISSLER, A. E. 1980. Auxin Transport Inhibitors .4. Evidence of a Common-Mode of Action for a Proposed Class of Auxin Transport Inhibitors - the Phytotropins. *Plant Physiology,* 66**,** 1190-1195.

KOGA, J. 1995. Structure and function of indolepyruvate decarboxylase, a key enzyme in indole-3-acetic acid biosynthesis. *Biochim Biophys Acta,* 1249**,** 1-13.

KOGA, J., ADACHI, T. & HIDAKA, H. 1992. Purification and characterization of indolepyruvate decarboxylase. A novel enzyme for indole-3-acetic acid biosynthesis in Enterobacter cloacae. *J Biol Chem,* 267**,** 15823-8.

KRECEK, P., SKUPA, P., LIBUS, J., NARAMOTO, S., TEJOS, R., FRIML, J. & ZAZIMALOVA, E. 2009. The PIN-FORMED (PIN) protein family of auxin transporters. *Genome Biol,* 10**,** 249.

KRIECHBAUMER, V., BOTCHWAY, S. W. & HAWES, C. 2016. Localization and interactions between Arabidopsis auxin biosynthetic enzymes in the TAA/YUC-dependent pathway. *J Exp Bot,* 67**,** 4195-207.

LAST, R. L., BISSINGER, P. H., MAHONEY, D. J., RADWANSKI, E. R. & FINK, G. R. 1991. Tryptophan Mutants in Arabidopsis - the Consequences of Duplicated Tryptophan Synthase Beta Genes. *Plant Cell,* 3**,** 345-358.

LAST, R. L. & FINK, G. R. 1988. Tryptophan-Requiring Mutants of the Plant Arabidopsis thaliana. *Science,* 240**,** 305-10.

LEHMANN, T., HOFFMANN, M., HENTRICH, M. & POLLMANN, S. 2010. Indole-3-acetamide-dependent auxin biosynthesis: a widely distributed way of indole-3-acetic acid production? *Eur J Cell Biol,* 89**,** 895-905.

LING, Y., ALSHAREEF, S., BUTT, H., LOZANO-JUSTE, J., LI, L., GALAL, A. A., MOUSTAFA, A., MOMIN, A. A., TASHKANDI, M., RICHARDSON, D. N., FUJII, H., AROLD, S., RODRIGUEZ, P. L., DUQUE, P. & MAHFOUZ, M. M. 2017. Pre-mRNA splicing repression triggers abiotic stress signaling in plants. *Plant J,* 89**,** 291-309.

LIU, J., MEHDI, S., TOPPING, J., FRIML, J. & LINDSEY, K. 2013a. Interaction of PLS and PIN and hormonal crosstalk in Arabidopsis root development. *Front Plant Sci,* 4**,** 75.

LIU, J., MEHDI, S., TOPPING, J., FRIML, J. & LINDSEY, K. 2013b. Interaction of PLS and PIN and hormonal crosstalk in Arabidopsis root development. *Frontiers in plant science,* 4**,** 75-75.

LIU, J., MOORE, S., CHEN, C. & LINDSEY, K. 2017. Crosstalk Complexities between Auxin, Cytokinin, and Ethylene in Arabidopsis Root Development: From Experiments to Systems Modeling, and Back Again. *Mol Plant,* 10**,** 1480-1496.

LIU, J. L., MEHDI, S., TOPPING, J., TARKOWSKI, P. & LINDSEY, K. 2010. Modelling and experimental analysis of hormonal crosstalk in Arabidopsis. *Molecular Systems Biology,* 6.

LJUNG, K. 2013. Auxin metabolism and homeostasis during plant development. *Development,* 140**,** 943-50.

LJUNG, K., HULL, A. K., CELENZA, J., YAMADA, M., ESTELLE, M., NORMANLY, J. & SANDBERG, G. 2005. Sites and regulation of auxin biosynthesis in Arabidopsis roots. *Plant Cell,* 17**,** 1090-104.

MAKAROVA, O. V., MAKAROV, E. M. & LUHRMANN, R. 2001. The 65 and 110 kDa SR-related proteins of the U4/U6.U5 tri-snRNP are essential for the assembly of mature spliceosomes. *EMBO J,* 20**,** 2553-63.

MANLEY, L. J., MA, D. & LEVINE, S. S. 2016. Monitoring Error Rates In Illumina Sequencing. *J Biomol Tech,* 27**,** 125-128.

MANO, Y. & NEMOTO, K. 2012. The pathway of auxin biosynthesis in plants. *Journal of Experimental Botany,* 63**,** 2853-2872.

MANO, Y., NEMOTO, K., SUZUKI, M., SEKI, H., FUJII, I. & MURANAKA, T. 2010. The AMI1 gene family: indole-3-acetamide hydrolase functions in auxin biosynthesis in plants. *Journal of Experimental Botany,* 61**,** 25-32.

MASHIGUCHI, K., TANAKA, K., SAKAI, T., SUGAWARA, S., KAWAIDE, H., NATSUME, M., HANADA, A., YAENO, T., SHIRASU, K., YAO, H., MCSTEEN, P., ZHAO, Y. D., HAYASHI, K., KAMIYA, Y. & KASAHARA, H. 2011. The main auxin biosynthesis pathway in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 18512-18517.

MIKKELSEN, M. D., NAUR, P. & HALKIER, B. A. 2004. Arabidopsis mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant J,* 37**,** 770-7.

MOORE, S., ZHANG, X., MUDGE, A., ROWE, J. H., TOPPING, J. F., LIU, J. & LINDSEY, K. 2015. Spatiotemporal modelling of hormonal crosstalk explains the level and patterning of hormones and gene expression in Arabidopsis thaliana wild-type and mutant roots. *New Phytol,* 207**,** 1110-22.

MUDGE, A. 2016. *The role of the POLARIS peptide in ethylene signalling and root development in Arabidopsis thaliana.* Doctoral Thesis, Durham University.

NILSEN, T. W. & GRAVELEY, B. R. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature,* 463**,** 457-63.

NORMANLY, J. 2010. Approaching Cellular and Molecular Resolution of Auxin Biosynthesis and Metabolism. *Cold Spring Harbor Perspectives in Biology,* 2.

POLLMANN, S., MULLER, A., PIOTROWSKI, M. & WEILER, E. W. 2002. Occurrence and formation of indole-3-acetamide in Arabidopsis thaliana. *Planta,* 216**,** 155-61.

RAMAKERS, C., RUIJTER, J. M., DEPREZ, R. H. & MOORMAN, A. F. 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett,* 339**,** 62-6.

REED, R. C., BRADY, S. R. & MUDAY, G. K. 1998. Inhibition of auxin movement from the shoot into the root inhibits lateral root development in arabidopsis. *Plant Physiology,* 118**,** 1369-1378.

RODRIGUEZ, F. I., ESCH, J. J., HALL, A. E., BINDER, B. M., SCHALLER, G. E. & BLEECKER, A. B. 1999. A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. *Science,* 283**,** 996-8.

SMART, C. M., SCOFIELD, S. R., BEVAN, M. W. & DYER, T. A. 1991. Delayed Leaf Senescence in Tobacco Plants Transformed with tmr, a Gene for Cytokinin Production in Agrobacterium. *Plant Cell,* 3**,** 647-656.

SOMERVILLE, C. & KOORNNEEF, M. 2002. A fortunate choice: the history of Arabidopsis as a model plant. *Nat Rev Genet,* 3**,** 883-9.

SONGSTAD, D. D., DELUCA, V., BRISSON, N., KURZ, W. G. W. & NESSLER, C. L. 1990. High-Levels of Tryptamine Accumulation in Transgenic Tobacco Expressing Tryptophan Decarboxylase. *Plant Physiology,* 94**,** 1410-1413.

STEPANOVA, A. N., HOYT, J. M., HAMILTON, A. A. & ALONSO, J. M. 2005. A link between ethylene and auxin uncovered by the characterization of two root-specific ethylene-insensitive mutants in Arabidopsis. *Plant Cell,* 17**,** 2230-2242.

STEPANOVA, A. N., ROBERTSON-HOYT, J., YUN, J., BENAVENTE, L. M., XIE, D. Y., DOLEZAL, K., SCHLERETH, A., JURGENS, G. & ALONSO, J. M. 2008. TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell,* 133**,** 177-191.

STEPANOVA, A. N., YUN, J., LIKHACHEVA, A. V. & ALONSO, J. M. 2007. Multilevel interactions between ethylene and auxin in Arabidopsis roots. *Plant Cell,* 19**,** 2169-85.

SUGAWARA, S., HISHIYAMA, S., JIKUMARU, Y., HANADA, A., NISHIMURA, T., KOSHIBA, T., ZHAO, Y., KAMIYA, Y. & KASAHARA, H. 2009. Biochemical analyses of indole-3-acetaldoxime-dependent auxin biosynthesis in Arabidopsis. *Proc Natl Acad Sci U S A,* 106**,** 5430-5.

SWARUP, R., PERRY, P., HAGENBEEK, D., VAN DER STRAETEN, D., BEEMSTER, G. T., SANDBERG, G., BHALERAO, R., LJUNG, K. & BENNETT, M. J. 2007. Ethylene upregulates auxin biosynthesis in Arabidopsis seedlings to enhance inhibition of root cell elongation. *Plant Cell,* 19**,** 2186-96.

TAO, Y., FERRER, J. L., LJUNG, K., POJER, F., HONG, F. X., LONG, J. A., LI, L., MORENO, J. E., BOWMAN, M. E., IVANS, L. J., CHENG, Y. F., LIM, J., ZHAO, Y. D., BALLARE, C. L., SANDBERG, G., NOEL, J. P. & CHORY, J. 2008. Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. *Cell,* 133**,** 164-176.

TIVENDALE, N. D., DAVIES, N. W., MOLESWORTH, P. P., DAVIDSON, S. E., SMITH, J. A., LOWE, E. K., REID, J. B. & ROSS, J. J. 2010. Reassessing the role of N-hydroxytryptamine in auxin biosynthesis. *Plant Physiol,* 154**,** 1957-65.

TIVENDALE, N. D., ROSS, J. J. & COHEN, J. D. 2014. The shifting paradigms of auxin biosynthesis. *Trends Plant Sci,* 19**,** 44-51.

TOPPING, J. F., AGYEMAN, F., HENRICOT, B. & LINDSEY, K. 1994. Identification of molecular markers of embryogenesis in Arabidopsis thaliana by promoter trapping. *Plant J,* 5**,** 895-903.

TOPPING, J. F. & LINDSEY, K. 1997. Promoter trap markers differentiate structural and positional components of polar development in Arabidopsis. *Plant Cell,* 9**,** 1713-25.

WINTER, A. 1966. A hypothetical route for the biogenisis of IAA. *Planta,* 71**,** 229-39.

WOODWARD, A. W. & BARTEL, B. 2005. Auxin: Regulation, action, and interaction. *Annals of Botany,* 95**,** 707-735.

YAMADA, T., PALM, C. J., BROOKS, B. & KOSUGE, T. 1985. Nucleotide sequences of the Pseudomonas savastanoi indoleacetic acid genes show homology with Agrobacterium tumefaciens T-DNA. *Proc Natl Acad Sci U S A,* 82**,** 6522-6.

ZHANG, R., CALIXTO, C. P. G., MARQUEZ, Y., VENHUIZEN, P., TZIOUTZIOU, N. A., GUO, W., SPENSLEY, M., ENTIZNE, J. C., LEWANDOWSKA, D., TEN HAVE, S., FREI DIT FREY, N., HIRT, H., JAMES, A. B., NIMMO, H. G., BARTA, A., KALYNA, M. & BROWN, J. W. S. 2017. A high quality Arabidopsis transcriptome for accurate transcript-level analysis of alternative splicing. *Nucleic Acids Res,* 45**,** 5061-5073.

ZHANG, R., WANG, B., JIAN, O. Y., LI, J. Y. & WANG, Y. H. 2008. Arabidopsis indole synthase, a homolog of tryptophan synthase alpha, is an enzyme involved in the Trp-independent indole-containing metabolite biosynthesis. *Journal of Integrative Plant Biology,* 50**,** 1070-1077.

ZHAO, Y. D., CHRISTENSEN, S. K., FANKHAUSER, C., CASHMAN, J. R., COHEN, J. D., WEIGEL, D. & CHORY, J. 2001. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science,* 291**,** 306-309.

ZHENG, Z., GUO, Y., NOVAK, O., DAI, X., ZHAO, Y., LJUNG, K., NOEL, J. P. & CHORY, J. 2013. Coordination of auxin and ethylene biosynthesis by the aminotransferase VAS1. *Nat Chem Biol,* 9**,** 244-6.

Appendices

Appendix 1 List of enriched gene ontology terms for up regulated genes in *pls* mutant.

term_ID	representative	description	Abs Log10Pvalue
GO:0042221	response to chemical	response to chemical	9.0809
GO:0018298	protein-chromophore linkage	protein-chromophore linkage	8.4437
GO:0010033	response to chemical	response to organic substance	8.4437
GO:0050896	response to stimulus	response to stimulus	8.4437
GO:0009719	response to chemical	response to endogenous stimulus	7.7959
GO:0009725	response to chemical	response to hormone	7.2366
GO:0006950	response to chemical	response to stress	7.0132
GO:0071369	response to chemical	cellular response to ethylene stimulus	6.0706
GO:0000160	response to chemical	phosphorelay signal transduction system	5.5376
GO:0009628	response to chemical	response to abiotic stimulus	5.5229
GO:0009768	photosynthesis, light harvesting in photosystem I	photosynthesis, light harvesting in photosystem I	5.4437
GO:0044710	secondary metabolism	single-organism metabolic process	5.1024
GO:0009607	response to chemical	response to biotic stimulus	4.9208
GO:0009611	response to chemical	response to wounding	4.6576
GO:0051707	response to chemical	response to other organism	4.6383
GO:0019748	secondary metabolism	secondary metabolic process	4.6383
GO:1901700	response to chemical	response to oxygen-containing compound	4.6021
GO:0010200	response to chemical	response to chitin	4.301
GO:0044699	single-organism process	single-organism process	4.301
GO:0072359	circulatory system development	circulatory system development	4.301
GO:0001944	circulatory system development	vasculature development	4.301
GO:0009605	response to chemical	response to external stimulus	4.1612
GO:0036294	response to chemical	cellular response to decreased oxygen levels	4.0757
GO:0015979	photosynthesis	photosynthesis	4.0269
GO:0009645	response to chemical	response to low light intensity stimulus	3.8861
GO:1901698	response to chemical	response to nitrogen compound	3.8239
GO:0006952	response to chemical	defense response	3.699

GO:0009642	response to chemical	response to light intensity	3.4202
GO:0006955	response to chemical	immune response	3.2147
GO:0002376	immune system process	immune system process	2.6778
GO:0006979	response to chemical	response to oxidative stress	2.5229
GO:0044550	secondary metabolism	secondary metabolite biosynthetic process	2.5229
GO:0009269	response to chemical	response to desiccation	2.4318
GO:0055114	secondary metabolism	oxidation-reduction process	2.2596
GO:0010114	response to chemical	response to red light	2.2291
GO:0051704	multi-organism process	multi-organism process	2.1739
GO:0009651	response to chemical	response to salt stress	2.0809
GO:0070482	response to chemical	response to oxygen levels	2.0605
GO:0006091	generation of precursor metabolites and energy	generation of precursor metabolites and energy	1.8861
GO:1901362	secondary metabolism	organic cyclic compound biosynthetic process	1.699
GO:0001101	response to chemical	response to acid chemical	1.5528
GO:0009056	catabolism	catabolic process	1.4949
GO:0044711	secondary metabolism	single-organism biosynthetic process	1.4559
GO:0090487	secondary metabolism	secondary metabolite catabolic process	1.4089
GO:0009407	secondary metabolism	toxin catabolic process	1.4089
GO:0042545	cell wall modification	cell wall modification	1.4089
GO:0019438	secondary metabolism	aromatic compound biosynthetic process	1.3872
GO:0009409	response to chemical	response to cold	1.284
GO:0008152	metabolism	metabolic process	1.2596
GO:0009746	response to chemical	response to hexose	1.2218
GO:0051716	response to chemical	cellular response to stimulus	1.1805
GO:0010035	response to chemical	response to inorganic substance	1.1612
GO:1902600	hydrogen ion transmembrane transport	hydrogen ion transmembrane transport	1.1192
GO:0042744	hydrogen peroxide catabolism	hydrogen peroxide catabolic process	1.1135
GO:0042886	hydrogen ion transmembrane transport	amide transport	1.1135
GO:0044712	hydrogen peroxide catabolism	single-organism catabolic process	1.0362
GO:0007154	cell communication	cell communication	1.0044
GO:0034654	secondary metabolism	nucleobase-containing compound biosynthetic process	0.9586
GO:0023052	signaling	signaling	0.9208

GO:0018130	secondary metabolism	heterocycle biosynthetic process	0.9208
GO:0009168	secondary metabolism	purine ribonucleoside monophosphate biosynthetic process	0.9208
GO:0006355	secondary metabolism	regulation of transcription, DNA-templated	0.8539
GO:0009664	cell wall modification	plant-type cell wall organization	0.8539
GO:0046129	secondary metabolism	purine ribonucleoside biosynthetic process	0.7959
GO:0044763	secondary metabolism	single-organism cellular process	0.7696
GO:0007568	circulatory system development	aging	0.7696
GO:0006818	hydrogen ion transmembrane transport	hydrogen transport	0.7696
GO:0015833	hydrogen ion transmembrane transport	peptide transport	0.7696
GO:0006629	secondary metabolism	lipid metabolic process	0.6778
GO:0022900	photosynthesis, light harvesting in photosystem I	electron transport chain	0.6778
GO:0071395	response to chemical	cellular response to jasmonic acid stimulus	0.5686
GO:0080167	response to chemical	response to karrikin	0.4685
GO:0071554	cell wall organization or biogenesis	cell wall organization or biogenesis	0.4559
GO:0000272	hydrogen peroxide catabolism	polysaccharide catabolic process	0.4559
GO:0015698	hydrogen ion transmembrane transport	inorganic anion transport	0.4437
GO:0072593	reactive oxygen species metabolism	reactive oxygen species metabolic process	0.4318
GO:0034220	hydrogen ion transmembrane transport	ion transmembrane transport	0.3979
GO:0005975	carbohydrate metabolism	carbohydrate metabolic process	0.3372
GO:0016042	hydrogen peroxide catabolism	lipid catabolic process	0.3279
GO:0006811	hydrogen ion transmembrane transport	ion transport	0.301

Appendix 2 List of enriched gene ontology terms for down regulated genes in pls mutant.

term_ID	representative	description	Abs Log10pvalue
GO:0019761	sulfur compound biosynthesis	glucosinolate biosynthetic process	7.6021
GO:0044710	secondary metabolism	single-organism metabolic process	7.3188
GO:0044272	sulfur compound biosynthesis	sulfur compound biosynthetic process	7.0269
GO:0050896	response to stimulus	response to stimulus	7.0269
GO:0006950	response to stress	response to stress	6.7696
GO:0019748	secondary metabolism	secondary metabolic process	6.1487
GO:0055114	sulfur compound biosynthesis	oxidation-reduction process	6.1135
GO:0009628	response to stress	response to abiotic stimulus	5.301
GO:0044711	sulfur compound biosynthesis	single-organism biosynthetic process	5.301
GO:0006790	sulfur compound metabolism	sulfur compound metabolic process	4.4202
GO:0044699	single-organism process	single-organism process	4.3979
GO:1901700	response to stress	response to oxygen-containing compound	4.2924
GO:0009625	response to stress	response to insect	4.0458
GO:0001101	response to stress	response to acid chemical	3.3565
GO:0042221	response to stress	response to chemical	3.3372
GO:0006082	sulfur compound biosynthesis	organic acid metabolic process	3.3372
GO:0009607	response to stress	response to biotic stimulus	3.0269
GO:0080167	response to stress	response to karrikin	2.9208
GO:0051707	response to stress	response to other organism	2.8239
GO:0046394	sulfur compound biosynthesis	carboxylic acid biosynthetic process	2.8239
GO:0009409	response to stress	response to cold	2.7447
GO:0009605	response to stress	response to external stimulus	2.5376
GO:0009611	response to stress	response to wounding	2.4437
GO:0009414	response to stress	response to water deprivation	2.4437
GO:0036293	response to stress	response to decreased oxygen levels	2.3979
GO:0070482	response to stress	response to oxygen levels	2.3872
GO:0010033	response to stress	response to organic substance	2.1805
GO:0008152	metabolism	metabolic process	2.1739
GO:0044281	sulfur compound biosynthesis	small molecule metabolic process	2.1739
GO:0009813	sulfur compound biosynthesis	flavonoid biosynthetic process	1.9208
GO:0042445	hormone metabolism	hormone metabolic process	1.8239

GO:0009812	flavonoid metabolism	flavonoid metabolic process	1.6778
GO:0006952	response to stress	defense response	1.585
GO:0080134	response to stress	regulation of response to stress	1.585
GO:0009411	response to stress	response to UV	1.585
GO:0044283	sulfur compound biosynthesis	small molecule biosynthetic process	1.585
GO:0032787	sulfur compound biosynthesis	monocarboxylic acid metabolic process	1.5686
GO:0065008	hormone metabolism	regulation of biological quality	1.5686
GO:0009725	response to stress	response to hormone	1.4318
GO:0051704	multi-organism process	multi-organism process	1.4318
GO:0010035	response to stress	response to inorganic substance	1.2596
GO:0009719	response to stress	response to endogenous stimulus	1.1487
GO:0048583	response to stress	regulation of response to stimulus	1.1367
GO:0033554	response to stress	cellular response to stress	1.1367
GO:1901605	sulfur compound biosynthesis	alpha-amino acid metabolic process	1.0223
GO:0010817	hormone metabolism	regulation of hormone levels	0.9586
GO:1901615	organic hydroxy compound metabolism	organic hydroxy compound metabolic process	0.6383
GO:0006811	ion transmembrane transport	ion transport	0.4089
GO:0034220	ion transmembrane transport	ion transmembrane transport	0.3872
GO:0006979	response to stress	response to oxidative stress	0.3665
GO:0048878	hormone metabolism	chemical homeostasis	0.3665
GO:0044763	secondary metabolism	single-organism cellular process	0.3188

Appendix 3 List of enriched gene ontology terms for up regulated genes in PLSox over expressor.

Gene Ontology analysis on Up-regulated genes in PLSox overexpressor. q<0.05

term_ID	representative	description	ABS log10pvalue
GO:0009698	phenylpropanoid metabolism	phenylpropanoid metabolic process	6.6778
GO:0044710	phenylpropanoid metabolism	single-organism metabolic process	6.2676
GO:0019748	phenylpropanoid metabolism	secondary metabolic process	5.7447
GO:0044699	single-organism process	single-organism process	5.3768
GO:0042744	hydrogen peroxide catabolism	hydrogen peroxide catabolic process	4.7447
GO:0006950	response to stress	response to stress	4.1487
GO:0050896	response to stimulus	response to stimulus	4.0605
GO:0055114	phenylpropanoid metabolism	oxidation-reduction process	3.6383
GO:0009409	response to stress	response to cold	3.4437
GO:0006970	response to stress	response to osmotic stress	3.3565
GO:0009628	response to stress	response to abiotic stimulus	3.0506
GO:0044711	phenylpropanoid metabolism	single-organism biosynthetic process	3.0506
GO:0071554	cell wall organization or biogenesis	cell wall organization or biogenesis	3.0506
GO:0071555	cell wall organization	cell wall organization	2.9208
GO:0019722	response to stress	calcium-mediated signaling	2.7696
GO:0009651	response to stress	response to salt stress	2.7447
GO:0006979	response to stress	response to oxidative stress	2.6778
GO:0071702	organic substance transport	organic substance transport	2.6778
GO:0019932	response to stress	second-messenger-mediated signaling	2.5229
GO:0006869	organic substance transport	lipid transport	2.5229
GO:0045229	cell wall organization	external encapsulating structure organization	2.5229
GO:0009806	phenylpropanoid metabolism	lignan metabolic process	2.3768
GO:0009807	phenylpropanoid metabolism	lignan biosynthetic process	2.3768
GO:0010876	organic substance transport	lipid localization	2.3768
GO:0010200	response to stress	response to chitin	2.3098
GO:0072593	reactive oxygen species metabolism	reactive oxygen species metabolic process	2.284
GO:0042221	response to stress	response to chemical	2.2007

GO:0006629	phenylpropanoid metabolism	lipid metabolic process	1.8539
GO:0009267	response to stress	cellular response to starvation	1.8239
GO:0048767	root hair elongation	root hair elongation	1.7696
GO:1901698	response to stress	response to nitrogen compound	1.5086
GO:0071496	response to stress	cellular response to external stimulus	1.4437
GO:0072330	phenylpropanoid metabolism	monocarboxylic acid biosynthetic process	1.4089
GO:1901700	response to stress	response to oxygen-containing compound	1.3872
GO:0002213	response to stress	defense response to insect	1.3872
GO:0042343	phenylpropanoid metabolism	indole glucosinolate metabolic process	1.3872
GO:0034220	organic substance transport	ion transmembrane transport	1.3872
GO:0016143	S-glycoside metabolism	S-glycoside metabolic process	1.3872
GO:0009611	response to stress	response to wounding	1.3768
GO:0000038	phenylpropanoid metabolism	very long-chain fatty acid metabolic process	1.3768
GO:0044763	phenylpropanoid metabolism	single-organism cellular process	1.3768
GO:0016126	phenylpropanoid metabolism	sterol biosynthetic process	1.3768
GO:0008610	phenylpropanoid metabolism	lipid biosynthetic process	1.3768
GO:0006810	organic substance transport	transport	1.3768
GO:1901135	carbohydrate derivative metabolism	carbohydrate derivative metabolic process	1.3768
GO:0051179	localization	localization	1.1871
GO:0033036	organic substance transport	macromolecule localization	1.1549
GO:0006811	organic substance transport	ion transport	1.1427
GO:0040007	growth	growth	1.0809
GO:0022622	root hair elongation	root system development	1.0809
GO:1902578	organic substance transport	single-organism localization	1.041
GO:0071456	response to stress	cellular response to hypoxia	0.9586
GO:0009056	catabolism	catabolic process	0.9586
GO:0044248	hydrogen peroxide catabolism	cellular catabolic process	0.9586
GO:0010035	response to stress	response to inorganic substance	0.9208
GO:0090627	root hair elongation	plant epidermal cell differentiation	0.9208
GO:0042435	phenylpropanoid metabolism	indole-containing compound biosynthetic process	0.8861

GO:0006818	organic substance transport	hydrogen transport	0.8861
GO:0009741	response to stress	response to brassinosteroid	0.8239
GO:0042430	phenylpropanoid metabolism	indole-containing compound metabolic process	0.8239
GO:0055085	organic substance transport	transmembrane transport	0.8239
GO:0090662	organic substance transport	ATP hydrolysis coupled transmembrane transport	0.8239
GO:1901657	S-glycoside metabolism	glycosyl compound metabolic process	0.8239
GO:0010033	response to stress	response to organic substance	0.7959
GO:0021700	root hair elongation	developmental maturation	0.7696
GO:0009719	response to stress	response to endogenous stimulus	0.7447
GO:0032787	phenylpropanoid metabolism	monocarboxylic acid metabolic process	0.7212
GO:0009828	cell wall organization	plant-type cell wall loosening	0.6198
GO:0044085	cell wall organization	cellular component biogenesis	0.6021
GO:0006082	phenylpropanoid metabolism	organic acid metabolic process	0.585
GO:0042742	response to stress	defense response to bacterium	0.5528
GO:0009605	response to stress	response to external stimulus	0.5376
GO:0015698	organic substance transport	inorganic anion transport	0.4685
GO:0042254	cell wall organization	ribosome biogenesis	0.4559
GO:0009414	response to stress	response to water deprivation	0.3979
GO:0044281	phenylpropanoid metabolism	small molecule metabolic process	0.3565
GO:2001057	reactive nitrogen species metabolism	reactive nitrogen species metabolic process	0.3279
GO:0005976	polysaccharide metabolism	polysaccharide metabolic process	0.3188

Appendix 4 List of enriched gene ontology terms for down regulated genes in PLSox over expressor.

term_ID	**representative**	**description**	**Abs Log10Pvalue**
GO:0090304	RNA processing	nucleic acid metabolic process	17.8539
GO:0006996	organelle organization	organelle organization	17.7696
GO:0051276	organelle organization	chromosome organization	17.5086
GO:0046483	RNA processing	heterocycle metabolic process	17.3979
GO:0006139	RNA processing	nucleobase-containing compound metabolic process	17.2518

GO:1901360	organic cyclic compound metabolism	organic cyclic compound metabolic process	17.1367
GO:0006725	RNA processing	cellular aromatic compound metabolic process	16.8539
GO:0009987	cellular process	cellular process	16.8239
GO:0006396	RNA processing	RNA processing	15.5376
GO:0006397	RNA processing	mRNA processing	15.0132
GO:0044699	single-organism process	single-organism process	15
GO:0000003	reproduction	reproduction	13.3468
GO:0016070	RNA processing	RNA metabolic process	13.3468
GO:0009791	post-embryonic development	post-embryonic development	13.284
GO:0016071	RNA processing	mRNA metabolic process	13.284
GO:0061458	post-embryonic development	reproductive system development	13.0088
GO:0050789	negative regulation of biological process	regulation of biological process	12.8239
GO:0022414	reproductive process	reproductive process	12.6198
GO:0065007	biological regulation	biological regulation	12.1427
GO:0071840	cellular component organization or biogenesis	cellular component organization or biogenesis	12
GO:0006325	organelle organization	chromatin organization	10.3468
GO:0006259	RNA processing	DNA metabolic process	10.0362
GO:0008152	metabolism	metabolic process	9.7212
GO:0071704	organic substance metabolism	organic substance metabolic process	9.7212
GO:0044763	movement of cell or subcellular component	single-organism cellular process	9.6383
GO:0008380	RNA processing	RNA splicing	9.5086
GO:0032502	developmental process	developmental process	8.8539
GO:0016569	organelle organization	covalent chromatin modification	8.7959
GO:0044237	RNA processing	cellular metabolic process	8.7696
GO:0048856	post-embryonic development	anatomical structure development	8.7447
GO:0044238	organic substance metabolism	primary metabolic process	8.7212
GO:0043170	organic cyclic compound metabolism	macromolecule metabolic process	8.4685
GO:0043933	organelle organization	macromolecular complex subunit organization	8.1427
GO:0032501	multicellular organismal process	multicellular organismal process	7.6383
GO:0033554	cellular response to stress	cellular response to stress	7.3872
GO:0044260	RNA processing	cellular macromolecule metabolic process	7.3872

GO:0048519	negative regulation of biological process	negative regulation of biological process	7.2757
GO:0019222	negative regulation of biological process	regulation of metabolic process	7.1487
GO:0006807	nitrogen compound metabolism	nitrogen compound metabolic process	6.7959
GO:0034641	RNA processing	cellular nitrogen compound metabolic process	6.6383
GO:0010605	negative regulation of biological process	negative regulation of macromolecule metabolic process	6.5229
GO:0050896	response to stimulus	response to stimulus	6.4318
GO:0006950	cellular response to stress	response to stress	6.3665
GO:0010468	negative regulation of biological process	regulation of gene expression	6.0605
GO:0006928	movement of cell or subcellular component	movement of cell or subcellular component	5.8861
GO:0051716	cellular response to stress	cellular response to stimulus	5.8539
GO:0006075	movement of cell or subcellular component	(1->3)-beta-D-glucan biosynthetic process	5.1938
GO:0000819	organelle organization	sister chromatid segregation	5.0706
GO:0015979	photosynthesis	photosynthesis	4.9586
GO:0018130	RNA processing	heterocycle biosynthetic process	4.7447
GO:0010467	RNA processing	gene expression	4.6778
GO:1901362	RNA processing	organic cyclic compound biosynthetic process	4.6383
GO:0040029	negative regulation of biological process	regulation of gene expression, epigenetic	4.5229
GO:0019438	RNA processing	aromatic compound biosynthetic process	4.4202
GO:0051128	organelle organization	regulation of cellular component organization	4.3665
GO:1901699	cellular response to stress	cellular response to nitrogen compound	4.2007
GO:0034654	RNA processing	nucleobase-containing compound biosynthetic process	4.0605
GO:0048367	post-embryonic development	shoot system development	3.9208
GO:0048518	negative regulation of biological process	positive regulation of biological process	3.8861
GO:0007059	movement of cell or subcellular component	chromosome segregation	3.7959
GO:0010033	cellular response to stress	response to organic substance	3.6383
GO:1901700	cellular response to stress	response to oxygen-containing compound	3.6021

GO:0009416	cellular response to stress	response to light stimulus	3.6021
GO:0007018	movement of cell or subcellular component	microtubule-based movement	3.3872
GO:0007049	movement of cell or subcellular component	cell cycle	3.2757
GO:0000280	organelle organization	nuclear division	3.2518
GO:0010228	reproductive process	vegetative to reproductive phase transition of meristem	3.2518
GO:0042221	cellular response to stress	response to chemical	3.2076
GO:0009628	cellular response to stress	response to abiotic stimulus	2.9586
GO:0048285	organelle organization	organelle fission	2.8861
GO:0006310	RNA processing	DNA recombination	2.7959
GO:0051640	organelle localization	organelle localization	2.7447
GO:0051239	post-embryonic development	regulation of multicellular organismal process	2.699
GO:0018205	RNA processing	peptidyl-lysine modification	2.5229
GO:0044728	RNA processing	DNA methylation or demethylation	2.5229
GO:0030029	movement of cell or subcellular component	actin filament-based process	2.4949
GO:0006304	RNA processing	DNA modification	2.4685
GO:0030048	movement of cell or subcellular component	actin filament-based movement	2.4559
GO:0010608	negative regulation of biological process	posttranscriptional regulation of gene expression	2.3188
GO:0007389	post-embryonic development	pattern specification process	2.0706
GO:1901698	cellular response to stress	response to nitrogen compound	2.0362
GO:0033993	cellular response to stress	response to lipid	2.0177
GO:0048511	rhythmic process	rhythmic process	2
GO:0009743	cellular response to stress	response to carbohydrate	1.9586
GO:0009657	organelle organization	plastid organization	1.9208
GO:0007623	circadian rhythm	circadian rhythm	1.9208
GO:0048507	post-embryonic development	meristem development	1.8861
GO:0005986	movement of cell or subcellular component	sucrose biosynthetic process	1.8861
GO:0009605	cellular response to stress	response to external stimulus	1.7959
GO:0001101	cellular response to stress	response to acid chemical	1.7959
GO:0019253	movement of cell or subcellular component	reductive pentose-phosphate cycle	1.7959

GO:0016051	movement of cell or subcellular component	carbohydrate biosynthetic process	1.7959
GO:0019685	movement of cell or subcellular component	photosynthesis, dark reaction	1.7447
GO:0044249	RNA processing	cellular biosynthetic process	1.699
GO:1901576	RNA processing	organic substance biosynthetic process	1.699
GO:0003002	post-embryonic development	regionalization	1.6778
GO:0002376	immune system process	immune system process	1.6576
GO:0048522	negative regulation of biological process	positive regulation of cellular process	1.6576
GO:0009902	organelle localization	chloroplast relocation	1.6576
GO:0006403	organelle localization	RNA localization	1.6576
GO:0097305	cellular response to stress	response to alcohol	1.6576
GO:0000018	RNA processing	regulation of DNA recombination	1.6576
GO:0051028	organelle localization	mRNA transport	1.6383
GO:0006260	RNA processing	DNA replication	1.6383
GO:0007017	movement of cell or subcellular component	microtubule-based process	1.6198
GO:0043414	RNA processing	macromolecule methylation	1.6021
GO:0002252	immune effector process	immune effector process	1.5686
GO:0051644	organelle localization	plastid localization	1.5686
GO:0043038	RNA processing	amino acid activation	1.5686
GO:0006418	RNA processing	tRNA aminoacylation for protein translation	1.5686
GO:0015977	carbon fixation	carbon fixation	1.5376
GO:0009737	cellular response to stress	response to abscisic acid	1.5086
GO:0043412	RNA processing	macromolecule modification	1.5086
GO:0010072	reproductive process	primary shoot apical meristem specification	1.4815
GO:0051179	localization	localization	1.4685
GO:0009058	biosynthesis	biosynthetic process	1.4685
GO:0048583	cellular response to stress	regulation of response to stimulus	1.4559
GO:0044764	multi-organism cellular process	multi-organism cellular process	1.4089
GO:0044710	movement of cell or subcellular component	single-organism metabolic process	1.4089
GO:0006298	cellular response to stress	mismatch repair	1.3979
GO:0042742	cellular response to stress	defense response to bacterium	1.3979
GO:0015994	movement of cell or subcellular component	chlorophyll metabolic process	1.3979

GO:0015995	movement of cell or subcellular component	chlorophyll biosynthetic process	1.3372
GO:0090066	organelle organization	regulation of anatomical structure size	1.3279
GO:0032535	organelle organization	regulation of cellular component size	1.3279
GO:0009607	cellular response to stress	response to biotic stimulus	1.2596
GO:0034645	RNA processing	cellular macromolecule biosynthetic process	1.2596
GO:0034285	cellular response to stress	response to disaccharide	1.1871
GO:0051234	organelle localization	establishment of localization	1.1612
GO:0006302	cellular response to stress	double-strand break repair	1.1612
GO:0044262	movement of cell or subcellular component	cellular carbohydrate metabolic process	1.1612
GO:0009059	RNA processing	macromolecule biosynthetic process	1.1367
GO:0080135	cellular response to stress	regulation of cellular response to stress	1.1308
GO:0080188	RNA processing	RNA-directed DNA methylation	1.1192
GO:0033043	organelle organization	regulation of organelle organization	1.1135
GO:0051301	movement of cell or subcellular component	cell division	1.1135
GO:0006366	RNA processing	transcription from RNA polymerase II promoter	1.0655
GO:0033013	RNA processing	tetrapyrrole metabolic process	1.041
GO:0040034	post-embryonic development	regulation of development, heterochronic	1.0223
GO:0009719	cellular response to stress	response to endogenous stimulus	1.0223
GO:0006261	RNA processing	DNA-dependent DNA replication	1.0223
GO:0019752	RNA processing	carboxylic acid metabolic process	1.0223
GO:0044711	movement of cell or subcellular component	single-organism biosynthetic process	1.0088
GO:0016310	RNA processing	phosphorylation	1.0044
GO:0052545	organelle localization	callose localization	0.9208
GO:0051641	organelle localization	cellular localization	0.9208
GO:0032101	cellular response to stress	regulation of response to external stimulus	0.9208
GO:0018193	RNA processing	peptidyl-amino acid modification	0.9208
GO:1901361	RNA processing	organic cyclic compound catabolic process	0.9208

GO:0034660	RNA processing	ncRNA metabolic process	0.9208
GO:0042440	carbon fixation	pigment metabolic process	0.9208
GO:0044093	negative regulation of biological process	positive regulation of molecular function	0.8861
GO:0017038	organelle localization	protein import	0.8861
GO:0051704	multi-organism process	multi-organism process	0.8861
GO:0048532	post-embryonic development	anatomical structure arrangement	0.8539
GO:0033037	organelle localization	polysaccharide localization	0.7959
GO:0048878	organelle organization	chemical homeostasis	0.7696
GO:0040007	growth	growth	0.7696
GO:0000725	cellular response to stress	recombinational repair	0.7696
GO:0008213	RNA processing	protein alkylation	0.7696
GO:0005982	movement of cell or subcellular component	starch metabolic process	0.7696
GO:0065008	negative regulation of biological process	regulation of biological quality	0.7447
GO:0007033	organelle organization	vacuole organization	0.7212
GO:0019439	RNA processing	aromatic compound catabolic process	0.7212
GO:0051168	organelle localization	nuclear export	0.699
GO:0006082	RNA processing	organic acid metabolic process	0.6778
GO:0006401	RNA processing	RNA catabolic process	0.6778
GO:0016116	movement of cell or subcellular component	carotenoid metabolic process	0.6778
GO:0099402	post-embryonic development	plant organ development	0.6383
GO:0006289	cellular response to stress	nucleotide-excision repair	0.6383
GO:0051052	RNA processing	regulation of DNA metabolic process	0.6383

Appendix 5 List of enriched gene ontology terms for up regulated genes in *mdf-1* mutant.

term_ID	representative	description	Absolute Log10pvalue
GO:0050896	response to stimulus	response to stimulus	35.3279
GO:0042221	response to chemical	response to chemical	34.7959
GO:0006950	response to chemical	response to stress	34.0458
GO:1901700	response to chemical	response to oxygen-containing compound	28.1308
GO:0044710	proton transport	single-organism metabolic process	23.6778
GO:0009628	response to chemical	response to abiotic stimulus	23.4815
GO:0001101	response to chemical	response to acid chemical	21.7212
GO:0055114	cellular respiration	oxidation-reduction process	19.7959
GO:0006952	response to chemical	defense response	19.585
GO:0044699	single-organism process	single-organism process	19.4815
GO:0010033	response to chemical	response to organic substance	19.1871
GO:0009719	response to chemical	response to endogenous stimulus	17.5376
GO:0006979	response to chemical	response to oxidative stress	17
GO:0010035	response to chemical	response to inorganic substance	16.8539
GO:0045333	cellular respiration	cellular respiration	16.6576
GO:0051707	response to chemical	response to other organism	16.4559
GO:0009607	response to chemical	response to biotic stimulus	15.3979
GO:0006970	response to chemical	response to osmotic stress	15.1427
GO:0009651	response to chemical	response to salt stress	13.9586
GO:0009605	response to chemical	response to external stimulus	13.8239
GO:0010243	response to chemical	response to organonitrogen compound	13.0605
GO:0080167	response to chemical	response to karrikin	12.6198
GO:0044281	cellular respiration	small molecule metabolic process	11.7696
GO:0051704	multi-organism process	multi-organism process	11.301
GO:0044712	cellular respiration	single-organism catabolic process	11.0809
GO:0009753	response to chemical	response to jasmonic acid	10.0969
GO:0009611	response to chemical	response to wounding	10.0315
GO:0006082	cellular respiration	organic acid metabolic process	9.8539
GO:0009723	response to chemical	response to ethylene	9.8239
GO:1901698	response to chemical	response to nitrogen compound	9.585
GO:0002376	immune system process	immune system process	9.5528
GO:0044763	proton transport	single-organism cellular process	9.5086
GO:0009056	catabolism	catabolic process	8.7959

GO:0097305	response to chemical	response to alcohol	8.3979
GO:0001666	response to chemical	response to hypoxia	8.0088
GO:0006955	response to chemical	immune response	7.7696
GO:0006091	generation of precursor metabolites and energy	generation of precursor metabolites and energy	7.7212
GO:0019748	cellular respiration	secondary metabolic process	7.3372
GO:0007568	aging	aging	7.2924
GO:0009404	cellular respiration	toxin metabolic process	7.1549
GO:0080134	response to chemical	regulation of response to stress	7.1024
GO:0009873	response to chemical	ethylene-activated signaling pathway	6.5528
GO:1901657	cellular respiration	glycosyl compound metabolic process	6.3665
GO:0009813	flavonoid biosynthesis	flavonoid biosynthetic process	6.1308
GO:0010038	response to chemical	response to metal ion	6.0809
GO:0042430	indole-containing compound metabolism	indole-containing compound metabolic process	5.7212
GO:0009812	flavonoid metabolism	flavonoid metabolic process	5.5086
GO:0046034	cellular respiration	ATP metabolic process	5.3188
GO:0090487	cellular respiration	secondary metabolite catabolic process	5.301
GO:0009407	cellular respiration	toxin catabolic process	5.301
GO:0055085	proton transport	transmembrane transport	5.301
GO:0010150	aging	leaf senescence	5.1367
GO:0010260	aging	animal organ senescence	5.1367
GO:0000160	response to chemical	phosphorelay signal transduction system	5.0969
GO:0012501	cellular respiration	programmed cell death	5.0862
GO:0032787	cellular respiration	monocarboxylic acid metabolic process	4.9208
GO:0008219	cellular respiration	cell death	4.8861
GO:0044711	cellular respiration	single-organism biosynthetic process	4.8539
GO:0048583	response to chemical	regulation of response to stimulus	4.7696
GO:0008152	metabolism	metabolic process	4.3565
GO:0009626	cellular respiration	plant-type hypersensitive response	4.0458
GO:0023052	signaling	signaling	4.0088
GO:0007154	cell communication	cell communication	4
GO:0034050	cellular respiration	host programmed cell death induced by symbiont	3.9586
GO:0009696	cellular respiration	salicylic acid metabolic process	3.9208
GO:1901135	carbohydrate derivative metabolism	carbohydrate derivative metabolic process	3.699
GO:0051716	response to chemical	cellular response to stimulus	3.6383
GO:0072593	reactive oxygen species metabolism	reactive oxygen species metabolic process	3.6021

GO:0034440	cellular respiration	lipid oxidation	3.5686
GO:0006811	proton transport	ion transport	3.5686
GO:0015992	proton transport	proton transport	3.4559
GO:0006818	proton transport	hydrogen transport	3.4559
GO:0072521	indole-containing compound metabolism	purine-containing compound metabolic process	3.3565
GO:0000154	indole-containing compound metabolism	rRNA modification	3.3372
GO:0009697	cellular respiration	salicylic acid biosynthetic process	3.3098
GO:0030258	cellular respiration	lipid modification	3.2518
GO:0044283	cellular respiration	small molecule biosynthetic process	3.2441
GO:1901605	cellular respiration	alpha-amino acid metabolic process	3.2218
GO:0044242	cellular respiration	cellular lipid catabolic process	3.2007
GO:0010193	response to chemical	response to ozone	3.1079
GO:0006855	response to chemical	drug transmembrane transport	3.0915
GO:1902578	proton transport	single-organism localization	3.0223
GO:0055086	cellular respiration	nucleobase-containing small molecule metabolic process	2.9208
GO:0006102	cellular respiration	isocitrate metabolic process	2.9208
GO:0006839	proton transport	mitochondrial transport	2.8861
GO:0046942	proton transport	carboxylic acid transport	2.8539
GO:0044419	interspecies interaction between organisms	interspecies interaction between organisms	2.8539
GO:0006790	sulfur compound metabolism	sulfur compound metabolic process	2.8539
GO:0005996	cellular respiration	monosaccharide metabolic process	2.7959
GO:0009853	photorespiration	photorespiration	2.7959
GO:0042493	response to chemical	response to drug	2.7696
GO:0009068	cellular respiration	aspartate family amino acid catabolic process	2.7696
GO:0042435	indole-containing compound metabolism	indole-containing compound biosynthetic process	2.7696
GO:0006810	proton transport	transport	2.6778
GO:0072350	cellular respiration	tricarboxylic acid metabolic process	2.6576
GO:0048513	aging	animal organ development	2.6383
GO:0019725	cellular respiration	cellular homeostasis	2.5376
GO:0006733	oxidoreduction coenzyme metabolism	oxidoreduction coenzyme metabolic process	2.5376
GO:0051179	localization	localization	2.5376
GO:0042181	cellular respiration	ketone biosynthetic process	2.4949
GO:0031407	cellular respiration	oxylipin metabolic process	2.4949
GO:0052314	cellular respiration	phytoalexin metabolic process	2.4559
GO:0009700	cellular respiration	indole phytoalexin biosynthetic process	2.4559

GO:0006560	cellular respiration	proline metabolic process	2.4559
GO:0042180	cellular respiration	cellular ketone metabolic process	2.4318
GO:0006787	cellular respiration	porphyrin-containing compound catabolic process	2.3872
GO:0033015	cellular respiration	tetrapyrrole catabolic process	2.3872
GO:0015748	proton transport	organophosphate ester transport	2.3872
GO:0051187	oxidoreduction coenzyme metabolism	cofactor catabolic process	2.3872
GO:0080024	indole-containing compound metabolism	indolebutyric acid metabolic process	2.3665
GO:0046487	cellular respiration	glyoxylate metabolic process	2.3468
GO:0072524	indole-containing compound metabolism	pyridine-containing compound metabolic process	2.2924
GO:0002213	response to chemical	defense response to insect	2.284
GO:0046149	cellular respiration	pigment catabolic process	2.284
GO:0042743	hydrogen peroxide metabolism	hydrogen peroxide metabolic process	2.284
GO:0006862	proton transport	nucleotide transport	2.1871
GO:0031408	cellular respiration	oxylipin biosynthetic process	2.1805
GO:0010498	cellular respiration	proteasomal protein catabolic process	2.1675
GO:0009694	cellular respiration	jasmonic acid metabolic process	2.1367
GO:0071695	aging	anatomical structure maturation	2.0969
GO:0009835	aging	fruit ripening	2.0969
GO:0042440	cellular respiration	pigment metabolic process	2.0506
GO:0009072	cellular respiration	aromatic amino acid family metabolic process	2.0269
GO:0016143	cellular respiration	S-glycoside metabolic process	2.0223
GO:0042343	cellular respiration	indole glucosinolate metabolic process	2.0223
GO:0051186	cofactor metabolism	cofactor metabolic process	2.0132
GO:0030162	cellular respiration	regulation of proteolysis	2
GO:0006097	cellular respiration	glyoxylate cycle	1.9586
GO:0043094	flavonoid biosynthesis	cellular metabolic compound salvage	1.9586
GO:0006022	cellular respiration	aminoglycan metabolic process	1.9208
GO:1990542	proton transport	mitochondrial transmembrane transport	1.8861
GO:0070588	proton transport	calcium ion transmembrane transport	1.8539
GO:0009744	response to chemical	response to sucrose	1.7959
GO:0046348	cellular respiration	amino sugar catabolic process	1.7959
GO:0046939	cellular respiration	nucleotide phosphorylation	1.7959

GO:0006121	cellular respiration	mitochondrial electron transport, succinate to ubiquinone	1.7959
GO:0044723	cellular respiration	single-organism carbohydrate metabolic process	1.7696
GO:0016053	cellular respiration	organic acid biosynthetic process	1.7696
GO:0009132	cellular respiration	nucleoside diphosphate metabolic process	1.6576
GO:0006561	cellular respiration	proline biosynthetic process	1.6576
GO:1901362	cellular respiration	organic cyclic compound biosynthetic process	1.6576
GO:0050789	cellular respiration	regulation of biological process	1.6198
GO:0009828	indole-containing compound metabolism	plant-type cell wall loosening	1.585
GO:0009695	cellular respiration	jasmonic acid biosynthetic process	1.5528
GO:0009987	cellular process	cellular process	1.5528

Appendix 6 List of enriched gene ontology terms for down regulated genes in *mdf-1* mutant.

term_ID	representative	description	Abs Log10pvalue
GO:0044699	single-organism process	single-organism process	25.1805
GO:0071554	cell wall organization or biogenesis	cell wall organization or biogenesis	13.4202
GO:0044763	secondary metabolism	single-organism cellular process	13.041
GO:0015979	photosynthesis	photosynthesis	11.1938
GO:0009888	tissue development	tissue development	10.6778
GO:0071555	cell wall organization	cell wall organization	9.9586
GO:0007167	enzyme linked receptor protein signaling pathway	enzyme linked receptor protein signaling pathway	9.6383
GO:0032502	developmental process	developmental process	9.1612
GO:0045229	cell wall organization	external encapsulating structure organization	9.1079
GO:0007166	enzyme linked receptor protein signaling pathway	cell surface receptor signaling pathway	8.4559
GO:0006928	movement of cell or subcellular component	movement of cell or subcellular component	8.0044
GO:0032501	multicellular organismal process	multicellular organismal process	7.9586
GO:0019684	photosynthesis, light reaction	photosynthesis, light reaction	7.585
GO:0044707	single-multicellular organism process	single-multicellular organism process	7.585
GO:0006468	protein phosphorylation	protein phosphorylation	7.4318
GO:0005976	protein phosphorylation	polysaccharide metabolic process	7.2518
GO:0044710	secondary metabolism	single-organism metabolic process	7.1675
GO:0090627	tissue development	plant epidermal cell differentiation	6.5528
GO:0005975	carbohydrate metabolism	carbohydrate metabolic process	6.5376
GO:0040008	enzyme linked receptor protein signaling pathway	regulation of growth	6.4949
GO:0007018	movement of cell or subcellular component	microtubule-based movement	5.9586
GO:0050794	enzyme linked receptor protein signaling pathway	regulation of cellular process	5.8861
GO:0007049	movement of cell or subcellular component	cell cycle	5.7212
GO:0007017	movement of cell or subcellular component	microtubule-based process	5.6778
GO:0051301	movement of cell or subcellular component	cell division	5.6383

GO:0009733	enzyme linked receptor protein signaling pathway	response to auxin	5.5229
GO:0010817	enzyme linked receptor protein signaling pathway	regulation of hormone levels	5.4318
GO:0019748	secondary metabolism	secondary metabolic process	5.3979
GO:0055114	secondary metabolism	oxidation-reduction process	5.0655
GO:0023052	signaling	signaling	5.0362
GO:0040007	growth	growth	5
GO:0022402	movement of cell or subcellular component	cell cycle process	4.9208
GO:0044711	secondary metabolism	single-organism biosynthetic process	4.8861
GO:0009914	enzyme linked receptor protein signaling pathway	hormone transport	4.8239
GO:0065007	biological regulation	biological regulation	4.6576
GO:0009606	enzyme linked receptor protein signaling pathway	tropism	4.6021
GO:0007154	cell communication	cell communication	4.5528
GO:0006629	secondary metabolism	lipid metabolic process	4.2518
GO:0044550	secondary metabolism	secondary metabolite biosynthetic process	4
GO:0009629	enzyme linked receptor protein signaling pathway	response to gravity	3.9586
GO:0006720	movement of cell or subcellular component	isoprenoid metabolic process	3.8239
GO:0018298	protein phosphorylation	protein-chromophore linkage	3.7212
GO:0042744	hydrogen peroxide catabolism	hydrogen peroxide catabolic process	3.2757
GO:0006793	protein phosphorylation	phosphorus metabolic process	3.1871
GO:0050896	response to stimulus	response to stimulus	2.9586
GO:0015995	secondary metabolism	chlorophyll biosynthetic process	2.699
GO:0006820	anion transport	anion transport	2.6778
GO:0045168	tissue development	cell-cell signaling involved in cell fate commitment	2.6576
GO:0009719	enzyme linked receptor protein signaling pathway	response to endogenous stimulus	2.3979
GO:0030104	enzyme linked receptor protein signaling pathway	water homeostasis	2.3565
GO:1901362	secondary metabolism	organic cyclic compound biosynthetic process	2.2757
GO:1905177	tissue development	tracheary element differentiation	2.1675
GO:0022403	cell cycle phase	cell cycle phase	2.1675
GO:0044848	biological phase	biological phase	2.1675
GO:0008356	movement of cell or subcellular component	asymmetric cell division	2.1675

GO:0071840	cellular component organization or biogenesis	cellular component organization or biogenesis	2.1612
GO:0006091	generation of precursor metabolites and energy	generation of precursor metabolites and energy	2.1079
GO:0031425	protein phosphorylation	chloroplast RNA processing	2.0269
GO:1900865	protein phosphorylation	chloroplast RNA modification	2.0044
GO:0015849	anion transport	organic acid transport	2
GO:0019438	secondary metabolism	aromatic compound biosynthetic process	2
GO:0015706	anion transport	nitrate transport	1.9586
GO:0007043	cell wall organization	cell-cell junction assembly	1.9586
GO:0000003	reproduction	reproduction	1.9208
GO:0044093	enzyme linked receptor protein signaling pathway	positive regulation of molecular function	1.8861
GO:0051049	anion transport	regulation of transport	1.8539
GO:0006811	anion transport	ion transport	1.7959
GO:0071705	anion transport	nitrogen compound transport	1.7959
GO:0016572	protein phosphorylation	histone phosphorylation	1.7447
GO:0008284	enzyme linked receptor protein signaling pathway	positive regulation of cell proliferation	1.7212
GO:0033013	protein phosphorylation	tetrapyrrole metabolic process	1.7212
GO:0042793	protein phosphorylation	transcription from plastid promoter	1.699
GO:0009625	enzyme linked receptor protein signaling pathway	response to insect	1.6778
GO:0016128	movement of cell or subcellular component	phytosteroid metabolic process	1.6778
GO:0030029	movement of cell or subcellular component	actin filament-based process	1.6576
GO:0009987	cellular process	cellular process	1.6383
GO:0044272	secondary metabolism	sulfur compound biosynthetic process	1.6383
GO:0007059	movement of cell or subcellular component	chromosome segregation	1.6198
GO:1902578	anion transport	single-organism localization	1.585
GO:0018130	secondary metabolism	heterocycle biosynthetic process	1.585
GO:0015977	secondary metabolism	carbon fixation	1.5229
GO:0008283	cell proliferation	cell proliferation	1.5229
GO:0070589	cell wall organization	cellular component macromolecule biosynthetic process	1.4949
GO:0045962	tissue development	positive regulation of development, heterochronic	1.4089

Appendix 7 List of enriched gene ontology terms for up regulated genes in *mdf-2* mutant.

term_ID	representative	description	Abs Log10pvalue
GO:0080134	response to stress	regulation of response to stress	6.1249
GO:0080167	response to stress	response to karrikin	4.1249
GO:0009409	response to stress	response to cold	3.8539
GO:0010243	response to stress	response to organonitrogen compound	18.8861
GO:0006955	response to stress	immune response	13.8861
GO:0001666	response to stress	response to hypoxia	7.6198
GO:0009414	response to stress	response to water deprivation	10.4815
GO:0001101	response to stress	response to acid chemical	22.9208
GO:0009744	response to stress	response to sucrose	1.6198
GO:0006952	response to stress	defense response	46.7212
GO:0033037	response to stress	polysaccharide localization	1.699
GO:0010286	response to stress	heat acclimation	1.6778
GO:0045229	response to stress	external encapsulating structure organization	1.4949
GO:2000022	response to stress	regulation of jasmonic acid mediated signaling pathway	2.1367
GO:0009605	response to stress	response to external stimulus	28.6383
GO:0009628	response to stress	response to abiotic stimulus	13.3872
GO:0000160	response to stress	phosphorelay signal transduction system	5.7447
GO:0052545	response to stress	callose localization	1.8539
GO:0052386	response to stress	cell wall thickening	2.4437
GO:0009723	response to stress	response to ethylene	9.0862
GO:0006970	response to stress	response to osmotic stress	8.1549
GO:0010646	response to stress	regulation of cell communication	1.4437
GO:0009719	response to stress	response to endogenous stimulus	17.3565
GO:0010035	response to stress	response to inorganic substance	9.4089
GO:0006979	response to stress	response to oxidative stress	14.0269
GO:0010033	response to stress	response to organic substance	20.5086
GO:0009625	response to stress	response to insect	2.1367
GO:0052544	response to stress	defense response by callose deposition in cell wall	3.1938
GO:0009743	response to stress	response to carbohydrate	1.3872
GO:0002239	response to stress	response to oomycetes	2.6576
GO:0048583	response to stress	regulation of response to stimulus	4.0757
GO:0033554	response to stress	cellular response to stress	2.3372
GO:0042493	response to stress	response to drug	2.3279
GO:0006855	response to stress	drug transmembrane transport	2.585

GO:0009607	response to stress	response to biotic stimulus	33.1079
GO:0009873	response to stress	ethylene-activated signaling pathway	7.1367
GO:0042221	response to stress	response to chemical	28.2924
GO:0002229	response to stress	defense response to oomycetes	2.301
GO:0002213	response to stress	defense response to insect	2.2676
GO:0051707	response to stress	response to other organism	33.8861
GO:1901698	response to stress	response to nitrogen compound	14.4559
GO:0051716	response to stress	cellular response to stimulus	8.2757
GO:0009651	response to stress	response to salt stress	6.9208
GO:0009611	response to stress	response to wounding	13.6383
GO:1901700	response to stress	response to oxygen-containing compound	27.8239
GO:0006950	response to stress	response to stress	49.1367
GO:0097305	response to stress	response to alcohol	7.8861
GO:0009646	response to stress	response to absence of light	2.3565
GO:0023051	response to stress	regulation of signaling	1.3372
GO:0042542	response to stress	response to hydrogen peroxide	3.1739
GO:0010193	response to stress	response to ozone	2.8239
GO:0044763	secondary metabolism	single-organism cellular process	4.9586
GO:0009407	secondary metabolism	toxin catabolic process	2.8861
GO:0090487	secondary metabolism	secondary metabolite catabolic process	2.8861
GO:0044711	secondary metabolism	single-organism biosynthetic process	3.2757
GO:0006629	secondary metabolism	lipid metabolic process	1.5376
GO:2000762	secondary metabolism	regulation of phenylpropanoid metabolic process	1.4318
GO:1901362	secondary metabolism	organic cyclic compound biosynthetic process	1.5528
GO:0044550	secondary metabolism	secondary metabolite biosynthetic process	9.2291
GO:0009700	secondary metabolism	indole phytoalexin biosynthetic process	3.4815
GO:0055114	secondary metabolism	oxidation-reduction process	15.585
GO:0044281	secondary metabolism	small molecule metabolic process	1.8539
GO:0019748	secondary metabolism	secondary metabolic process	11.7447
GO:0044710	secondary metabolism	single-organism metabolic process	14.4949
GO:0006040	programmed cell death	amino sugar metabolic process	2.284

GO:0006022	programmed cell death	aminoglycan metabolic process	3.0458
GO:1901657	programmed cell death	glycosyl compound metabolic process	1.7212
GO:0008202	programmed cell death	steroid metabolic process	1.5686
GO:0031408	programmed cell death	oxylipin biosynthetic process	2.8239
GO:0009626	programmed cell death	plant-type hypersensitive response	5.7447
GO:0031407	programmed cell death	oxylipin metabolic process	3.2518
GO:0006694	programmed cell death	steroid biosynthetic process	1.4437
GO:0050789	programmed cell death	regulation of biological process	3.699
GO:0044272	programmed cell death	sulfur compound biosynthetic process	1.2518
GO:0034050	programmed cell death	host programmed cell death induced by symbiont	5.6576
GO:0030258	programmed cell death	lipid modification	2.8861
GO:0016101	programmed cell death	diterpenoid metabolic process	1.3279
GO:0034440	programmed cell death	lipid oxidation	3.5229
GO:0008219	programmed cell death	cell death	8.5086
GO:0016143	programmed cell death	S-glycoside metabolic process	4.9208
GO:0009851	programmed cell death	auxin biosynthetic process	1.3279
GO:0006355	programmed cell death	regulation of transcription, DNA-templated	1.4685
GO:0006874	programmed cell death	cellular calcium ion homeostasis	3.2676
GO:0045333	programmed cell death	cellular respiration	1.5229
GO:0045927	programmed cell death	positive regulation of growth	1.5229
GO:0048518	programmed cell death	positive regulation of biological process	2.1549
GO:0042445	programmed cell death	hormone metabolic process	1.5376
GO:0012501	programmed cell death	programmed cell death	8
GO:0016042	cellular catabolism	lipid catabolic process	2.699
GO:0044242	cellular catabolism	cellular lipid catabolic process	3.5376
GO:0046149	cellular catabolism	pigment catabolic process	2.8239

GO:0009074	cellular catabolism	aromatic amino acid family catabolic process	2.1612
GO:0033015	cellular catabolism	tetrapyrrole catabolic process	2.4949
GO:1901136	cellular catabolism	carbohydrate derivative catabolic process	3.7212
GO:0046348	cellular catabolism	amino sugar catabolic process	3.3665
GO:0042545	cellular catabolism	cell wall modification	1.6576
GO:0010411	cellular catabolism	xyloglucan metabolic process	1.8239
GO:0044248	cellular catabolism	cellular catabolic process	10.2757
GO:0015996	cellular catabolism	chlorophyll catabolic process	2.9208
GO:0016998	cellular catabolism	cell wall macromolecule catabolic process	3.1675
GO:0042744	cellular catabolism	hydrogen peroxide catabolic process	3.3279
GO:0042787	cellular catabolism	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	2.0315
GO:0051187	cellular catabolism	cofactor catabolic process	2.4949
GO:0050896	response to stimulus	response to stimulus	42.7447
GO:0009299	indole-containing compound metabolism	mRNA transcription	1.5229
GO:0010337	indole-containing compound metabolism	regulation of salicylic acid metabolic process	2.2518
GO:0006082	indole-containing compound metabolism	organic acid metabolic process	3.0706
GO:0006560	indole-containing compound metabolism	proline metabolic process	2.8539
GO:0042435	indole-containing compound metabolism	indole-containing compound biosynthetic process	4.5528
GO:0006576	indole-containing compound metabolism	cellular biogenic amine metabolic process	1.5086
GO:0009696	indole-containing compound metabolism	salicylic acid metabolic process	2
GO:0009695	indole-containing compound metabolism	jasmonic acid biosynthetic process	2.0605
GO:0042180	indole-containing compound metabolism	cellular ketone metabolic process	1.699
GO:0009694	indole-containing compound metabolism	jasmonic acid metabolic process	2.0655

GO:0009064	indole-containing compound metabolism	glutamine family amino acid metabolic process	2.5229
GO:0042430	indole-containing compound metabolism	indole-containing compound metabolic process	8.8539
GO:0009072	indole-containing compound metabolism	aromatic amino acid family metabolic process	1.8861
GO:0032787	indole-containing compound metabolism	monocarboxylic acid metabolic process	2.1135
GO:0016053	indole-containing compound metabolism	organic acid biosynthetic process	1.3872
GO:0010260	aging	animal organ senescence	6.5376
GO:0048317	aging	seed morphogenesis	2.4685
GO:0001763	aging	morphogenesis of a branching structure	1.5528
GO:0009956	aging	radial pattern formation	1.2924
GO:0010150	aging	leaf senescence	6.5376
GO:0048513	aging	animal organ development	4.8239
GO:0007568	aging	aging	9.8861
GO:0051704	multi-organism process	multi-organism process	24.0915
GO:0055085	calcium ion transmembrane transport	transmembrane transport	2.7447
GO:0015749	calcium ion transmembrane transport	monosaccharide transport	1.2757
GO:0046323	calcium ion transmembrane transport	glucose import	1.4318
GO:0046942	calcium ion transmembrane transport	carboxylic acid transport	1.7959
GO:0070588	calcium ion transmembrane transport	calcium ion transmembrane transport	2.2676
GO:0006820	calcium ion transmembrane transport	anion transport	1.7696
GO:0043090	calcium ion transmembrane transport	amino acid import	1.7959
GO:0006869	calcium ion transmembrane transport	lipid transport	1.2366

GO:0015833	calcium ion transmembrane transport	peptide transport	1.4437
GO:0030001	calcium ion transmembrane transport	metal ion transport	1.3979
GO:0015849	calcium ion transmembrane transport	organic acid transport	2.0915
GO:0006811	calcium ion transmembrane transport	ion transport	1.7212
GO:0044699	single-organism process	single-organism process	19.2218
GO:0002376	immune system process	immune system process	16.7447
GO:0009056	catabolism	catabolic process	9.9586
GO:0030162	protein phosphorylation	regulation of proteolysis	2.8239
GO:0016310	protein phosphorylation	phosphorylation	2.1024
GO:0045732	protein phosphorylation	positive regulation of protein catabolic process	1.5229
GO:0006468	protein phosphorylation	protein phosphorylation	3.4685
GO:0007154	cell communication	cell communication	8.5229
GO:0023052	signaling	signaling	8.4202
GO:0006012	cellular carbohydrate metabolism	galactose metabolic process	1.5528
GO:0044262	cellular carbohydrate metabolism	cellular carbohydrate metabolic process	1.5686
GO:0005991	cellular carbohydrate metabolism	trehalose metabolic process	1.3872
GO:0072593	reactive oxygen species metabolism	reactive oxygen species metabolic process	4.4685
GO:0065007	biological regulation	biological regulation	2.8239
GO:0008152	metabolism	metabolic process	2.7696
GO:0044419	interspecies interaction between organisms	interspecies interaction between organisms	2.699
GO:0005975	carbohydrate metabolism	carbohydrate metabolic process	2.1612
GO:0006790	sulfur compound metabolism	sulfur compound metabolic process	1.4685
GO:0071554	cell wall organization or biogenesis	cell wall organization or biogenesis	1.3565
GO:2001057	reactive nitrogen species metabolism	reactive nitrogen species metabolic process	1.3468

Appendix 8 List of enriched gene ontology terms for down regulated genes in *mdf-2* mutant.

term_ID	representative	description	Abs Log10pvalue
GO:0000003	reproduction	reproduction	1.3279
GO:0005976	polysaccharide metabolism	polysaccharide metabolic process	7.7447
GO:0044723	polysaccharide metabolism	single-organism carbohydrate metabolic process	1.5686
GO:0016051	polysaccharide metabolism	carbohydrate biosynthetic process	3.6021
GO:0016310	polysaccharide metabolism	phosphorylation	1.6383
GO:0044264	polysaccharide metabolism	cellular polysaccharide metabolic process	5.699
GO:0044262	polysaccharide metabolism	cellular carbohydrate metabolic process	5.3188
GO:0018298	polysaccharide metabolism	protein-chromophore linkage	2.1427
GO:0016572	polysaccharide metabolism	histone phosphorylation	2.4685
GO:0006468	polysaccharide metabolism	protein phosphorylation	3.699
GO:0008152	metabolism	metabolic process	1.2676
GO:0009733	response to auxin	response to auxin	10.6021
GO:0071241	response to auxin	cellular response to inorganic substance	1.6383
GO:0010218	response to auxin	response to far red light	2.3098
GO:0071281	response to auxin	cellular response to iron ion	1.7959
GO:0006355	response to auxin	regulation of transcription, DNA-templated	1.5528
GO:0050794	response to auxin	regulation of cellular process	3.9586
GO:0007267	response to auxin	cell-cell signaling	1.7959
GO:0009630	response to auxin	gravitropism	4.7212
GO:0009629	response to auxin	response to gravity	3.9586
GO:0071731	response to auxin	response to nitric oxide	1.2441
GO:0010583	response to auxin	response to cyclopentenone	6.0862
GO:0009637	response to auxin	response to blue light	2.1249
GO:0009606	response to auxin	tropism	4.8539
GO:1902170	response to auxin	cellular response to reactive nitrogen species	1.284
GO:0007167	response to auxin	enzyme linked receptor protein signaling pathway	7.6383
GO:0007166	response to auxin	cell surface receptor signaling pathway	5.301
GO:0007165	response to auxin	signal transduction	4.7959
GO:0010033	response to auxin	response to organic substance	1.9208

GO:0009740	response to auxin	gibberellic acid mediated signaling pathway	1.3098
GO:0009719	response to auxin	response to endogenous stimulus	2.8861
GO:0009314	response to auxin	response to radiation	2.0044
GO:0009888	tissue development	tissue development	11.2007
GO:0007389	tissue development	pattern specification process	2.8539
GO:0003002	tissue development	regionalization	2.4685
GO:0099402	tissue development	plant organ development	6.2291
GO:0048507	tissue development	meristem development	1.5376
GO:0090558	tissue development	plant epidermis development	9.0223
GO:0021700	tissue development	developmental maturation	2.699
GO:0022622	tissue development	root system development	8.1308
GO:0010089	tissue development	xylem development	3.7959
GO:0010087	tissue development	phloem or xylem histogenesis	3.3188
GO:0048469	tissue development	cell maturation	4.585
GO:0030855	tissue development	epithelial cell differentiation	2.7959
GO:0045165	tissue development	cell fate commitment	3.2076
GO:0010453	tissue development	regulation of cell fate commitment	3
GO:0048367	tissue development	shoot system development	1.4318
GO:0048364	tissue development	root development	8.2147
GO:0010374	tissue development	stomatal complex development	3.8539
GO:0009791	tissue development	post-embryonic development	1.2676
GO:0048646	tissue development	anatomical structure formation involved in morphogenesis	1.284
GO:0009886	tissue development	post-embryonic animal morphogenesis	1.6198
GO:0090627	tissue development	plant epidermal cell differentiation	6.8239
GO:1905177	tissue development	tracheary element differentiation	3.8239
GO:0010817	regulation of hormone levels	regulation of hormone levels	5.8861
GO:0051338	regulation of hormone levels	regulation of transferase activity	1.7212
GO:0008284	regulation of hormone levels	positive regulation of cell proliferation	2.2366
GO:0035266	regulation of hormone levels	meristem growth	2.8539
GO:0032147	regulation of hormone levels	activation of protein kinase activity	1.9208
GO:0040008	regulation of hormone levels	regulation of growth	5.8239
GO:0010252	regulation of hormone levels	auxin homeostasis	2.8861
GO:0009914	regulation of hormone levels	hormone transport	5.8539
GO:0023052	signaling	signaling	5.1079
GO:0032501	multicellular organismal process	multicellular organismal process	6.6778

GO:0032502	developmental process	developmental process	8.0269
GO:0040007	growth	growth	5.9208
GO:0044699	single-organism process	single-organism process	23.4685
GO:0050896	response to stimulus	response to stimulus	3.0088
GO:0065007	biological regulation	biological regulation	3.2291
GO:0071554	cell wall organization or biogenesis	cell wall organization or biogenesis	15.8539
GO:0071840	cellular component organization or biogenesis	cellular component organization or biogenesis	3.1739
GO:0042744	hydrogen peroxide catabolism	hydrogen peroxide catabolic process	3.9586
GO:0044273	hydrogen peroxide catabolism	sulfur compound catabolic process	1.8539
GO:0042546	cell wall biogenesis	cell wall biogenesis	11.8539
GO:0000280	cell wall biogenesis	nuclear division	5.3665
GO:0034329	cell wall biogenesis	cell junction assembly	2.4685
GO:0034330	cell wall biogenesis	cell junction organization	1.9586
GO:1903340	cell wall biogenesis	positive regulation of cell wall organization or biogenesis	2.2757
GO:1903338	cell wall biogenesis	regulation of cell wall organization or biogenesis	2.2366
GO:0044085	cell wall biogenesis	cellular component biogenesis	1.5686
GO:0030261	cell wall biogenesis	chromosome condensation	1.5686
GO:0048285	cell wall biogenesis	organelle fission	4.0655
GO:0009834	cell wall biogenesis	plant-type secondary cell wall biogenesis	8.7447
GO:0009828	cell wall biogenesis	plant-type cell wall loosening	1.7212
GO:0007010	cell wall biogenesis	cytoskeleton organization	3.2757
GO:0070589	cell wall biogenesis	cellular component macromolecule biosynthetic process	2.8239
GO:0045229	cell wall biogenesis	external encapsulating structure organization	8.6778
GO:0009812	flavonoid metabolism	flavonoid metabolic process	1.5229
GO:0009768	photosynthesis, light harvesting in photosystem I	photosynthesis, light harvesting in photosystem I	2.3468
GO:0007049	cell cycle	cell cycle	10.699
GO:0019762	cell cycle	glucosinolate catabolic process	2.0809
GO:0033384	cell cycle	geranyl diphosphate biosynthetic process	2.2291
GO:0033383	cell cycle	geranyl diphosphate metabolic process	2.2291
GO:0033386	cell cycle	geranylgeranyl diphosphate biosynthetic process	1.9586

GO:0033385	cell cycle	geranylgeranyl diphosphate metabolic process	1.9586
GO:0019742	cell cycle	pentacyclic triterpenoid metabolic process	1.8239
GO:0019745	cell cycle	pentacyclic triterpenoid biosynthetic process	1.8239
GO:0045337	cell cycle	farnesyl diphosphate biosynthetic process	2.2291
GO:0051302	cell cycle	regulation of cell division	1.3372
GO:0051301	cell cycle	cell division	10.1675
GO:0044550	cell cycle	secondary metabolite biosynthetic process	4.4089
GO:0008610	cell cycle	lipid biosynthetic process	3.5229
GO:0006720	cell cycle	isoprenoid metabolic process	4.6021
GO:0006722	cell cycle	triterpenoid metabolic process	2.3665
GO:0008299	cell cycle	isoprenoid biosynthetic process	4.6383
GO:0061640	cell cycle	cytoskeleton-dependent cytokinesis	3.4685
GO:0007059	cell cycle	chromosome segregation	3.0757
GO:0007017	cell cycle	microtubule-based process	5.6383
GO:0007018	cell cycle	microtubule-based movement	5.9586
GO:0022402	cell cycle	cell cycle process	9.5376
GO:0030029	cell cycle	actin filament-based process	1.699
GO:0006928	cell cycle	movement of cell or subcellular component	6.9586
GO:0030048	cell cycle	actin filament-based movement	1.6383
GO:0008356	cell cycle	asymmetric cell division	3.0555
GO:1901617	organic hydroxy compound biosynthesis	organic hydroxy compound biosynthetic process	2.3665
GO:0009813	organic hydroxy compound biosynthesis	flavonoid biosynthetic process	1.8861
GO:0015979	photosynthesis	photosynthesis	2.0969
GO:1901615	organic hydroxy compound metabolism	organic hydroxy compound metabolic process	2.3979
GO:0072593	reactive oxygen species metabolism	reactive oxygen species metabolic process	1.7696
GO:0007154	cell communication	cell communication	5.2441
GO:0005975	carbohydrate metabolism	carbohydrate metabolic process	5.6021
GO:0019748	secondary metabolism	secondary metabolic process	8.8239
GO:0044711	secondary metabolism	single-organism biosynthetic process	6
GO:0044710	secondary metabolism	single-organism metabolic process	7.9208
GO:0044763	secondary metabolism	single-organism cellular process	13.3098

GO:0055114	secondary metabolism	oxidation-reduction process	5.0605
GO:0019438	secondary metabolism	aromatic compound biosynthetic process	1.2924
GO:0006629	secondary metabolism	lipid metabolic process	4.699
GO:0046148	secondary metabolism	pigment biosynthetic process	2.699
GO:1901362	secondary metabolism	organic cyclic compound biosynthetic process	1.6778
GO:0042440	secondary metabolism	pigment metabolic process	1.7959
GO:0008283	cell proliferation	cell proliferation	1.2518
GO:0035825	reciprocal DNA recombination	reciprocal DNA recombination	2.8861
GO:0006865	amino acid transport	amino acid transport	2.8539
GO:0006820	amino acid transport	anion transport	1.3768
GO:0006869	amino acid transport	lipid transport	1.4437
GO:0010876	amino acid transport	lipid localization	1.5229

Appendix 9 List of enriched gene ontology terms for differentially spliced genes in *mdf-1* mutant.

term_ID	representative	description	Abs Log10Pvalue
GO:0000003	mRNA metabolism	nucleobase-containing compound metabolic process	9.2076
GO:0002376	mRNA metabolism	nucleic acid metabolic process	9.0555
GO:0007623	mRNA metabolism	mRNA metabolic process	7.2757
GO:0008152	mRNA metabolism	mRNA processing	6.9586
GO:0009987	mRNA metabolism	RNA metabolic process	6.7447
GO:0010050	mRNA metabolism	RNA processing	6.5528
GO:0044706	mRNA metabolism	RNA splicing	5.2441
GO:0099402	mRNA metabolism	macromolecule modification	5.0605
GO:0048510	mRNA metabolism	cellular macromolecule metabolic process	4.8861
GO:0048507	mRNA metabolism	phosphate-containing compound metabolic process	4.1675
GO:0022622	mRNA metabolism	DNA metabolic process	4.1367
GO:0009553	mRNA metabolism	tRNA metabolic process	3.5528
GO:0048367	mRNA metabolism	regulation of RNA splicing	3.1135
GO:0048364	mRNA metabolism	organic cyclic compound biosynthetic process	3.0315
GO:0009791	mRNA metabolism	heterocycle biosynthetic process	3
GO:0040034	mRNA metabolism	ncRNA metabolic process	2.9586
GO:0009888	mRNA metabolism	cellular nitrogen compound metabolic process	2.6778
GO:0061458	mRNA metabolism	peptidyl-lysine modification	2.6021
GO:0016071	mRNA metabolism	histone H3-K4 methylation	2.585
GO:0006306	mRNA metabolism	protein phosphorylation	2.3979
GO:0018193	mRNA metabolism	nucleobase-containing compound biosynthetic process	2.1805
GO:0006796	mRNA metabolism	aromatic compound biosynthetic process	2.0969
GO:0018205	mRNA metabolism	steroid metabolic process	2.041
GO:0016070	mRNA metabolism	macromolecule methylation	2.041
GO:0044260	mRNA metabolism	macromolecule glycosylation	1.9208
GO:0044275	mRNA metabolism	RNA 3'-end processing	1.9208
GO:0006139	mRNA metabolism	nucleobase-containing compound catabolic process	1.8861
GO:0019438	mRNA metabolism	peptidyl-amino acid modification	1.8539
GO:0046488	mRNA metabolism	steroid biosynthetic process	1.7696
GO:0006694	mRNA metabolism	protein autophosphorylation	1.7696
GO:0090407	mRNA metabolism	phosphatidylinositol metabolic process	1.6778

GO:0006259	mRNA metabolism	nucleic acid phosphodiester bond hydrolysis	1.6778
GO:0090304	mRNA metabolism	gene expression	1.4949
GO:0090305	mRNA metabolism	protein alkylation	1.4559
GO:0043484	mRNA metabolism	phospholipid metabolic process	1.3188
GO:0010467	mRNA metabolism	protein acetylation	1.3098
GO:0018130	mRNA metabolism	organophosphate biosynthetic process	1.301
GO:0051568	mRNA metabolism	cellular carbohydrate catabolic process	1.1549
GO:0008202	mRNA metabolism	glyceraldehyde-3-phosphate metabolic process	1.0044
GO:0043414	mRNA metabolism	DNA methylation	1
GO:0043413	regulation of response to stress	negative regulation of biological process	4.041
GO:0006644	regulation of response to stress	regulation of response to stress	3.3872
GO:0043412	regulation of response to stress	regulation of biological process	3.1249
GO:0008213	regulation of response to stress	response to oxygen radical	3.1135
GO:0006396	regulation of response to stress	regulation of gene expression	3.041
GO:0008380	regulation of response to stress	immune response	3.0269
GO:0006397	regulation of response to stress	negative regulation of metabolic process	2.9586
GO:0006399	regulation of response to stress	regulation of metabolic process	2.8861
GO:0019682	regulation of response to stress	defense response, incompatible interaction	2.8239
GO:1901362	regulation of response to stress	cellular response to DNA damage stimulus	2.7696
GO:0034641	regulation of response to stress	regulation of response to stimulus	2.3468
GO:0006473	regulation of response to stress	defense response to fungus, incompatible interaction	2.2218
GO:0031123	regulation of response to stress	response to light stimulus	2.0809
GO:0046777	regulation of response to stress	regulation of DNA metabolic process	2.0223
GO:0006468	regulation of response to stress	response to inorganic substance	1.8861
GO:0034660	regulation of response to stress	regulation of DNA-dependent DNA replication	1.7959
GO:0034655	regulation of response to stress	cell cycle DNA replication	1.7447
GO:0034654	regulation of response to stress	response to acid chemical	1.699

GO:0022414	regulation of response to stress	DNA-dependent DNA replication	1.6383
GO:0010228	regulation of response to stress	auxin metabolic process	1.6198
GO:0009845	regulation of response to stress	cellular metal ion homeostasis	1.585
GO:0009910	regulation of response to stress	response to oxygen-containing compound	1.5686
GO:0032501	regulation of response to stress	cellular response to stimulus	1.4685
GO:0032502	regulation of response to stress	DNA recombination	1.4559
GO:0044699	regulation of response to stress	cold acclimation	1.3372
GO:0048511	regulation of response to stress	regulation of hormone levels	1.2441
GO:0048519	regulation of response to stress	telomere maintenance	1.2366
GO:0006304	regulation of response to stress	anatomical structure homeostasis	1.1938
GO:0090066	regulation of response to stress	response to abiotic stimulus	1.1549
GO:0048518	regulation of response to stress	DNA modification	1.0555
GO:1901700	regulation of response to stress	regulation of response to biotic stimulus	1.0555
GO:0006310	regulation of response to stress	positive regulation of biological process	1
GO:0000305	regulation of response to stress	detection of external stimulus	1
GO:0051716	regulation of response to stress	DNA replication	1
GO:0006261	regulation of response to stress	regulation of anatomical structure size	0.9586
GO:0048583	regulation of response to stress	reciprocal DNA recombination	0.9586
GO:0009581	organelle organization	chromosome organization	4.4437
GO:0044786	organelle organization	organelle organization	4.4202
GO:0006875	organelle organization	chloroplast organization	3.4318
GO:0019222	organelle organization	macromolecular complex subunit organization	3.3768
GO:0009628	organelle organization	organelle fission	2.9208
GO:0050789	organelle organization	plastid organization	2.5376
GO:0009631	organelle organization	chromatin organization	2.4437
GO:0002831	organelle organization	protein polymerization	2.041
GO:0006260	organelle organization	protein-DNA complex subunit organization	1.6198
GO:0010035	organelle organization	regulation of organelle organization	1.6021
GO:0001101	organelle organization	cellular component assembly	1.5686

GO:0010468	organelle organization	cellular component disassembly	1.2924
GO:0035825	organelle organization	telomere organization	1.2366
GO:0090329	organelle organization	regulation of cellular component biogenesis	1.1675
GO:0009814	organelle organization	ribonucleoprotein complex subunit organization	1.1135
GO:0009817	cofactor metabolism	heterocycle metabolic process	10.1675
GO:0051052	cofactor metabolism	cellular aromatic compound metabolic process	8.3979
GO:0060249	cofactor metabolism	cellular metabolic process	5.5376
GO:0009892	cofactor metabolism	phosphorus metabolic process	4.0088
GO:0010817	cofactor metabolism	cofactor metabolic process	1.8539
GO:0009850	cell recognition	cell recognition	2.8539
GO:0006974	cell recognition	cell cycle process	2.6021
GO:0009416	cell recognition	amino acid activation	2.3372
GO:0006955	cell recognition	sister chromatid cohesion	2.2218
GO:0080134	cell recognition	chromosome segregation	2.0706
GO:0000723	cell recognition	cell cycle	2.0506
GO:0051179	cell recognition	cellular amino acid metabolic process	1.5528
GO:0065007	cell recognition	polyol metabolic process	1.3188
GO:0071840	cell recognition	actin filament-based process	1.2076
GO:0072511	cell recognition	cell death	1.1805
GO:0055085	cell recognition	cellular ketone metabolic process	1.1675
GO:0006810	cell recognition	pyrimidine-containing compound biosynthetic process	1.1487
GO:0006811	cell recognition	water-soluble vitamin metabolic process	1.0132
GO:0070838	cell recognition	programmed cell death	0.9586
GO:0016197	cell recognition	amine metabolic process	0.9586
GO:0042886	vegetative phase change	vegetative phase change	3.3468
GO:0030001	vegetative phase change	post-embryonic development	2.3768
GO:0032259	vegetative phase change	multi-multicellular organism process	1.8861
GO:1900673	vegetative phase change	regulation of timing of transition from vegetative to reproductive phase	1.8539
GO:0006996	vegetative phase change	regulation of development, heterochronic	1.8239
GO:0071824	vegetative phase change	plant organ development	1.6021
GO:0071826	vegetative phase change	meristem development	1.6021
GO:0051258	vegetative phase change	shoot system development	1.585

GO:0051276	vegetative phase change	reproductive system development	1.3665
GO:0006325	vegetative phase change	root development	1.3279
GO:0022607	vegetative phase change	root system development	1.3098
GO:0044087	vegetative phase change	tissue development	1.0706
GO:0009658	vegetative phase change	embryo sac development	1.0362
GO:0009657	organic cyclic compound metabolism	organic cyclic compound metabolic process	9.7696
GO:0048285	organic cyclic compound metabolism	macromolecule metabolic process	4.6576
GO:0043933	organic cyclic compound metabolism	carbohydrate derivative metabolic process	1.7959
GO:0033043	organic cyclic compound metabolism	carbohydrate metabolic process	1.2676
GO:0022411	divalent inorganic cation transport	divalent inorganic cation transport	3.5528
GO:0032200	divalent inorganic cation transport	divalent metal ion transport	3.1612
GO:1901615	divalent inorganic cation transport	metal ion transport	2.4949
GO:0070085	divalent inorganic cation transport	ion transport	2.1249
GO:0044723	divalent inorganic cation transport	transport	1.7447
GO:0044710	divalent inorganic cation transport	endosomal transport	1.585
GO:0044763	divalent inorganic cation transport	transmembrane transport	1.4318
GO:0006629	divalent inorganic cation transport	amide transport	1.0555
GO:0044281	cofactor biosynthesis	cofactor biosynthetic process	2.6383
GO:0005982	cofactor biosynthesis	glycoprotein metabolic process	2.5086
GO:0008037	cofactor biosynthesis	carbohydrate derivative biosynthetic process	2.3188
GO:0019751	cofactor biosynthesis	glycoprotein biosynthetic process	2.284
GO:0012501	cofactor biosynthesis	ethylene biosynthetic process	1.4318
GO:0042180	cofactor biosynthesis	cellular metabolic compound salvage	1.1024
GO:0043038	cofactor biosynthesis	liposaccharide metabolic process	0.9586
GO:0006767	glycosylation	single-organism cellular process	2.9586
GO:0006520	glycosylation	glycosylation	2.9208
GO:0007062	glycosylation	single-organism metabolic process	2.4559

GO:0007049	glycosylation	single-organism carbohydrate metabolic process	1.3279
GO:0007059	glycosylation	small molecule metabolic process	1.2518
GO:0008219	glycosylation	starch metabolic process	1.1308
GO:0009308	glycosylation	lipid metabolic process	1.0269
GO:0022402	organic substance metabolism	organic substance metabolic process	5.7212
GO:0072528	organic substance metabolism	primary metabolic process	5.0655
GO:0030029	reproductive process	reproductive process	2.7212
GO:0051188	reproductive process	vegetative to reproductive phase transition of meristem	2.0862
GO:0009100	reproductive process	negative regulation of flower development	1.699
GO:0009101	reproductive process	seed germination	1.4815
GO:1901137	metabolism	metabolic process	6.7447
GO:0043094	cellular process	cellular process	3.9208
GO:0009693	rhythmic process	rhythmic process	3.5686
GO:1903509	single-organism process	single-organism process	3.3468
GO:1901360	circadian rhythm	circadian rhythm	3.1805
GO:0043170	nitrogen compound metabolism	nitrogen compound metabolic process	2.7447
GO:1901135	reproduction	reproduction	2.6778
GO:0005975	immune system process	immune system process	2.6198
GO:0071704	biological regulation	biological regulation	1.8539
GO:0044238	localization	localization	1.8239
GO:0051186	methylation	methylation	1.8239
GO:0006793	multicellular organismal process	multicellular organismal process	1.7959
GO:0044237	olefin metabolism	olefin metabolic process	1.4318
GO:0046483	developmental process	developmental process	1.3768
GO:0006725	organic hydroxy compound metabolism	organic hydroxy compound metabolic process	1.3188
GO:0006807	cellular component organization or biogenesis	cellular component organization or biogenesis	1.0706

Appendix 10 Primer List

Target Gene	Primer Sequence (5’ to 3’)
AOO1 F	GAAGCGAGTAGTGAAGCGGT
AOO1 R	TTTACTTCAACCTCGCTCGCT
ATM1 A F	TCGTCGGTGAAGCCAGAGTT
ATM1 A R	CACCCTTCTCCCCGAGCAAA
ATM1 B F	ATCTCGTCGGTGAAGCCAGA
ATM1 B R	TAGCTTGGAGATGCGGTGGA
ATMKK9 F	TCTCCGCCGTACAGATTCAC
ATMKK9 R	TCCGCCGTCCATATACTCCA
ATMYB122 F	GCCGTGTTGTAGAGCAGAAG
ATMYB122 R	GAATGGCGTGGAGGTTGATG
ATMYB34 F	CTTCTTAACCGCGTCGCAAG
ATMYB34 R	GCAATGTGGAGGTCGGAGAA
AUX1 F	TCGGTCCAATCAATTCCGCT
AUX1 R	AGAACCCAAACCACCACGAA
CYP71A13 F	TGTAGGCCGATTTGACTGGA
CYP71A13 R	TAGTGTTGCATAGCATAACAAGGT
CYP79B2 F	CCGTTGGTTATGGTCGGTGA
CYP79B2 R	GGCGTCGTCTCATCTCACTT
CYP79B3 F	GGCGTTAGGTACGGCGATAA
CYP79B3 R	CACGTGTCTCACTTCCTGCT
IAA2 C F	CCTCCTACCAAAACTCAAATCGTT
IAA2 C R	CGTAGCTCACACTGTTGTTGTTCT
LAX1 F	TCTGCTTCGGAGAGTGTAGAA
LAX1 R	AACGAATACGGCAGCGTCAA
LAX2 F	CTGCGGTTTATTGGGCGTTT
LAX2 R	ACCGAACCCGAATCCAACAA
LAX3 F	CTCGGACTCGCTATGACCAC

LAX3 R	CAAACGCCCAATAAACCGCA
MYB51 F	TCTTCCTTCGCACCAACCTC
MYB51 R	CGTGGACGAAGAACACGTCA
PAD3 F	CAGTGGCGCTTCTCCAATAC
PAD3 R	TAGCACTTCCTCTGCTGCTT
PIN1 B F	TCGTTGCTTCTTATGCCGTT
PIN1 B R	AGAAGAGTTATGGGCAACGC
PIN1 C F	CGGCGGCTATGAGATTTGTCGTT
PIN1 C R	CGCGATCAACATCCCAAATATCAC
PIN2 B F	AACGAGTGGAGCAAGTGGAG
PIN2 B R	CGGTGGAAGAACCTCTGGTC
PIN2 C F	AATGCTGGTTGCTTTGCCTG
PIN2 C R	CCTTTGGGTCGTATCGCCTT
SOT16 F	CGTCCTGCTGTTTATGCGAA
SOT16 R	TTTCGCCGCACCAAATAACA
SUR1 F	TGTTTCTACCAGGGGATGCT
SUR1 R	CTTTTGGGCACACACATCCT
TAA1 A F	GTGGTCAATCTGGATCATGGTG
TAA1 A R	CAGCGTTACCAACAACACCG
TAA1 B F	TGGTCAATCTGGATCATGGTG
TAA1 B R	CCGTACACCTGTCACCCATC
TSB1 A F	TATGAAGGTGGTAGCCCAGC
TSB1 A R	GCTACTTGCCATACTTCGCA
TSB1 B F	TGATTTTGGCGATACGGCTG
TSB1 B R	ACCGGCATGGATAGGATCAG
TSB2 F	CGAACTACCGATTTCCTCCAC
TSB2 R	ATTTGGGGAGTTTGGAGGGT
UGT74B1 F	ACAATGGCGGAAACAACTCC
UGT74B1 R	GTGGTGGCGATTGTGACTTT
WEI2 F	GGGTTATTTGCAGGCTAGAGG
WEI2 R	TCCCTCTCTTGCTGGTTCCT

WEI7 F	CTTCAACCCAAGTCTGGCTC
WEI7 R	ACGGAAACTCTTGTGGGATT
YUC2 F	CCCATGTGGCTAAAGGGAGTG
YUC2 R	CAATCCGCTTTCGCCTTTCC
YUC8 F	CGCGGTTAAGATCGCACAAG
YUC8 R	CTTGAGCGTTTCGTGGGTTG